UNIVERSITY OF BIRMINGHAM

Understanding the molecular basis of γδ T cell receptor ligand recognition in cellular stress surveillance

---

by

Mahboob Salim

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

Supervisor: Professor Benjamin E. Willcox

Abstract

γδ T-cells are unconventional lymphocytes hypothesised to act at the interface between innate and adaptive immunity. Emerging evidence indicates γδ T-cells play important, non-redundant roles in lymphoid stress surveillance during infection and tumourigenesis, and γδ T-cell-focussed immunotherapy trials suggest their potential exploitation in cancer immunotherapy. However, the molecular basis of γδ TCR/ligand recognition is poorly understood. In this thesis I first focussed on ligand recognition by the LES TCR, which is derived from a Vδ2-negative T-cell and mediates TCR-dependent recognition of CMV-infected cells and tumour cell lines by binding to Endothelial Protein C Receptor (EPCR). After producing LES TCR in a conformationally correct form, I used mutagenesis to map the LES TCR binding site on EPCR. Importantly, EPCR was recognised independently of bound lipid, suggesting it acts as a stress ligand rather than an antigen presenting molecule, and highlighting the importance of TCR-extrinsic factors in recognition. Secondly, I determined an NMR structure of Skint-1, a selecting ligand for mouse skin-resident DETC γδ T-cells that play important roles in immunoregulation and tumour immunosurveillance. This emphasised structural features unique to Skint-1, and suggested interaction with a separate ligand, such as the DETC TCR. Collectively, these studies improve understanding of γδ T-cell recognition.

Acknowledgements

I would firstly like to thank my supervisor Professor Benjamin Willcox for providing me with the opportunity to undertake a PhD and for his encouragement and enormous enthusiasm throughout this study. I would also like to thank Dr Carrie Willcox for her supervision and guidance on SPR experiments and for educating me in all things related to γδ T cell immunology. I am extremely grateful for all the help, guidance and advice provided by Dr Fiyaz Mohammed, particularly on the structural biology aspects of this study. I am also grateful to Carrie and Fiyaz for providing constructive criticism and proof reading this thesis. I would also like to thank Dr Timothy Knowles for his great assistance with NMR experiments and guidance with processing data. I am very grateful to Professor Adrian Hayday and Dr Julie Déchanet-Merville for kindly providing reagents to facilitate this research. A special mention goes to Sarah Nicholls who despite putting up with all my rants during the stressful times of my PhD has always provided helpful advice and friendship. I would also like to thank past and present members of the Willcox and Moss laboratories for their help and advice throughput the duration of my thesis. Finally, I would like to thank all my friends (especially Saj) and family who have supported me for the duration of this PhD, especially my partner Donna and my brothers.

CONTENTS

List of Figures

Chapter 4

Chapter 5

Chapter 6

List of Tables

List of Abbreviations

APC	Activated protein C
β_2M	β2-microglobulin
BCR	B cell receptor
C	Constant
CD	Cluster of differentiation
CDR	Complementarity determining region
CLIP	Class II-associated invariant chain peptide
CMV	Cytomegalovirus
CTL	Cytotoxic T lymphocytes
D	Diversity
DC	Dendritic cells
DNA	deoxyribonucleic acid
DES	Drosophila expression system
DETC	Dendritic epidermal T cells
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EDC	N-ethyl-N5 (3-diethylaminopropyl)-carbodiimide
EPCR	Endothelial Protein C receptor
ER	Endoplasmic Reticulum
Fab	Fragment antigen binding
Fc	Fragment crystallisable
FPLC	Fast Protein Liquid Chromatography
HLA	Human Leukocyte Antigen
HBS	Hepes Buffered Saline
HRP	Horseradish peroxidase
HSQC	Heteronuclear Single Quantum Coherence
IEL	Intraepithelial lymphocyte
IFN	Interferon
Ig	Immunoglobulin
IgH	Immunoglobulin Heavy chain
IgL	Immunoglobulin Light chain
IgSF	Immunoglobulin Super Family
IL	Interleukin
IPP	Isopentenyl pyrophosphate
IPTG	Isopropyl-β-D-thiogalactopyranside
J	Joining
LB	Lysogeny broth
KGF	Keratinocyte Growth Factor
mAb	Monoclonal antibody
MAD	Multi-wavelength anomalous dispersion
MAIT	Mucosal associated invariant T cells
MHC	Major Histocompatibility Complex

MICA	MHC class I polypeptide related sequence A
MIR	Multiple isomorphous replacement
MT	Metallothionien
Ni-NTA	Nickel-Nitrilotriacetic Acid
NK	Natural Killer
NKR	Natural Killer receptor
NMR	Nuclear Magnetic resonance
PAMP	Pathogen associated molecular patterns
PRR	Pattern recognition receptors
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDVF	Polyvinylidene difluoride
PLC	Peptide Loading Complex
PMSF	Phenylmethylsulfonyl fluoride
RAG	Recombination activating genes
siRNA	Small interfering ribonucleic acid
SDM	Site Directed Mutagenesis
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide gel electrophoresis
Skint	Selection and upkeep of intraepithelial T-cells
SPR	Surface plasmon resonance
TAP	Transport associated with antigen processing
TBS	Tris buffered saline
TCA	Trichloroacetic acid
TCR	T cell receptor
TdT	Terminal deoxynucleotidyl transferase
TEMED	Tetramethylethylenediamine
TGF-β	Transforming growth factor β
TLR	Toll-like receptor
T_H	Helper T cell
T_{regs}	Regulatory T cells
V	Variable
Wt	Wild type

CHAPTER 1 –

GENERAL INTRODUCTION

Chapter 1: Introduction

1.1 Vertebrate adaptive immunity

The vertebrate immune system appears to have evolved mainly to fight infection and is composed of two major arms: the innate and adaptive immune systems. The innate immune system probably evolved as a mechanism to protect against infection with common pathogens. The primary activators of innate immunity are innate pattern recognition receptors (PRRs), which include toll-like receptors (TLRs) (Medzhitov *et al*, 1997). The TLRs on dendritic cells (DC) and monocytes sense pathogen-associated molecular patterns (PAMPs) (Yang *et al*, 1998) and initiate effector responses such as phagocytosis, cytokine production and activation of the adaptive immune system. The innate immune system also includes Natural Killer (NK) cells, which detect pathogen infected cells through activating and inhibitory receptors on the cell surface.

The second arm to the immune system is the adaptive system, which evolved approximately 350-400 million years ago in jawed vertebrates. While PRRs are crucial to efficient activation of the innate immune system, the adaptive immune system evolved to use somatic gene rearrangement to generate diverse repertoire antigen receptors capable of detecting foreign molecules. Cells expressing useful receptors are amplified by clonal selection, and can be committed to immunological memory to provide faster responses upon secondary infection. The cells mainly responsible for adaptive immunity are B lymphocytes and αβ T lymphocytes. Whereas B cells recognise antigen through their cell surface B cell receptor (BCR) T cells use a homologous receptor termed the T cell receptor (TCR) (Hedrick *et al*, 1984a; Hedrick *et al*, 1984b). Most T cells use an αβ TCR, but there is a subpopulation of T

lymphocytes which express a different antigen receptor on the cell surface, a γδ T cell receptor (TCR). γδ T cells also use somatic rearrangement to generate their TCR, but the antigens recognised by the γδ TCR remain poorly understood. Also, since γδ T cells express many innate NK-like receptors, this potentially places them at the border of innate and adaptive immunity. This thesis focuses on molecular recognition by γδ T cells.

1.2 Antigen receptors

Adaptive immune cells possess the ability to recognise foreign molecules by the expression of unique antigen receptors expressed on their surface. The interaction of receptor with specific antigen initiates an intracellular signalling cascade which results in a variety of effector responses. In B cells, antigen binding to the B cell receptor (BCR) or surface immunoglobulin (Ig) leads to cellular differentiation and production of secreted Ig, or antibodies (Ab), while antigen recognition by T cells through their (TCR) leads to cytokine secretion or cytotoxicity.

The antigen receptors expressed by B cells and T cells are multimeric proteins grouped within immunoglobulin superfamily (IgSF). Antibodies are composed of two Ig heavy chains (IgH) and two Ig light chains (IgL). Each IgH consists of four domains of 110 amino acids, while each IgL chain consists of two such domains. Each 110 amino acid domain is called an Ig domain, and folds into 7-9 β-strands that form two anti-parallel β-sheets, stabilised by a disulphide bond. Disulphide bonds link the IgL to the IgH, as well as linking the two IgH chains in each antibody. Antibodies can be cleaved with the enzyme papain to generate three fragments. Two contain the antigen binding site and are called the Fragment antigen binding (Fab), while the other fragment is called the Fragment crystallisable (Fc), as this fragment was found to crystallise easily.

The antigen binding region of the Fab is confined to the N-terminal Ig domains of IgH and IgL. These domains have high sequence variability and are termed Ig variable (V) domains. The Ig domains not involved in antigen binding are less variable and are classed as Ig constant (C) regions. The Fc region of the antibody contains two C domains and is important for interactions with Fc receptors on innate cells, and with the complement cascade, which lead to clearance of pathogens.

Sequence variability within the V regions of IgH and IgL is focussed in three hypervariable regions in each chain (Wu and Kabat, 1970). These hypervariable regions correspond to the loop regions on the structure of a folded Ig protein. The three IgH hypervariable loops and three IgL hypervariable loops determine the antigen specificity, partly by establishing surface complementary to the antigen (Wilson *et al*, 1991), and therefore these loops were designated complementarity-determining regions (CDRs). The sequences in between the CDRs in the linear sequence are less variable and form the core β strands that provide the structure to the Ig fold, and are described as framework regions.

The vast variability of foreign antigens might suggest an equally vast number of possible antigen receptors genes, and yet the genome cannot possibly separately encode antigen receptors for all possible antigens. Instead, a process of gene rearrangement has evolved that enables production of a highly diverse receptor repertoire from limited genetic material. Antigen receptor genes are recombined from individual gene segments, a process involving DNA rearrangement and described as somatic recombination (Hozumi and Tonegawa, 1976). This process requires V (D) J recombinase enzymes, especially the recombination activating genes RAG-1 and RAG-2. In B cells, IgH recombination involves the joining of any of 250-

1000 variable (V_H) gene segments to one of ten diversity (D_H) gene segments and to one of four joining (J_H) gene segments (Table 1.1). The rearranged sequence is then spliced to the constant domain in the mRNA. These potential 40,000 V(D)J combinations can be further diversified by the addition of non-template nucleotides at regions of V(D)J joins by the lymphoid specific enzyme terminal deoxynucleotidyl transferase (TdT). The repertoire is increased further by the pairing of a rearranged IgH with a rearranged IgL. All these mechanisms result in the production of diverse antibody specificities which can recognise a plethora of different antigens.

The structure of TCRs is similar to immunoglobulins (Hedrick *et al*, 1984b), and TCR domains are therefore included in the IgSF. However, TCRs, unlike Ig, exist only in a membrane bound form. The majority of T cells express a heterodimeric TCR that is made up of α and β chains, while a small percentage express a γδ TCR. Each TCR chain consists of two extracellular Ig domains, composed of a membrane-distal V domain and a membrane-proximal C domain, as well as a transmembrane domain and short cytoplasmic tail. The two chains are connected by a disulphide bond located near the transmembrane domain. Amino acids in the transmembrane domain of the TCR chains facilitate association with the CD3 signalling complex.

The membrane distal domains of the TCR contain high sequence variation and are classed as IgV domains, while the membrane proximal domains do not vary in sequence and are classed as IgC domains. Similar to Ig, the variable domains of TCR chains each have three hypervariable regions which are also defined as complementarity determining regions (CDR), and the three CDR loops of TCRα combine with the three CDR loops of TCRβ to form the

antigen binding site. Like Ig, TCRs use gene rearrangement mediated by RAG-1 and RAG-2 to generate diverse TCRs to recognise foreign antigen. The TCRα and β gene loci contain fewer V gene segments than IgH and IgL, but more J segments (Table 1.1). In addition, the TCRβ chain uses D gene segments as does IgH, and as with Ig, diversity can be increased by N-nucleotide addition (Davis and Bjorkman, 1988).

1.3 Antigen recognition by conventional αβ T cells

Unlike B cells, αβ T cells do not recognise intact antigen. Instead, as demonstrated initially by Zinkernagel and Doherty (Zinkernagel and Doherty, 1974), αβ T cells are able to recognise antigen in a way that is restricted by self-encoded molecules of the major histocompatibility complex (MHC), which are able to bind and present short peptide fragments derived from foreign antigens directly to T cells (reviewed in Townsend *et al*, 1989). In humans, the MHC locus is termed the Human Leukocyte Antigen (HLA) locus, while in mice it is referred to as the histocompatibility 2 (H-2) complex. The functional consequence of peptide-MHC binding depends on which type of MHC molecule is presenting antigen, and which class of T cell is activated.

1.3.1 Conventional αβ T cell subsets

Conventional T cells are classified by their expression of CD4 or CD8 coreceptors, and these cell types exhibit different effector functions. CD8 αβ T cells recognise peptides derived from intracellular pathogens presented by MHC class I molecules (HLA-A, B and C in humans), differentiate into cytotoxic T lymphocytes (CTLs), which are able to kill infected cells. CD4 αβ T cells recognise peptides generated from extracellular pathogens presented by-MHC class II molecules, and can be subdivided further into subpopulations of T helper cells, based on the

profile of cytokines they produce (Mucida *et al*, 2010). T_H1 cells develop in the presence of IL-12, and secrete IFN-γ and TNFα, which enhance the killing of NK and CTL, and the microbicidal properties of macrophages. T_H2 cells develop in the presence of IL-4, and secrete cytokines which enhance B cell activation, differentiation and antibody secretion. T_H17 cells are a recently described subset that develops in the presence of IL-6 and TGFβ and produces IL-17 and IL-22 upon activation (Korn *et al*, 2009). T_H17 cells are crucial for the recruitment of innate cells such as neutrophils to the site of infection. Another important subset of T cells is the regulatory T cells (Tregs) (Vignali *et al*, 2008). These can be subdivided into natural Treg which develop in the thymus, or induced Treg which differentiate following infection in the presence of TGFβ and IL-2. When Treg cells recognise peptide antigen they suppress the activation of other specific T cells via production of immunosuppressive cytokines such as IL-10 and TGF-β, and also cell-cell contact-dependent mechanisms, so preventing autoimmune reactions or dampening down the immune response to prevent immunopathology.

1.3.2 Molecular mechanisms underlying αβ T cell recognition

Experiments by Townsend and colleagues proved that T cells recognise small protein fragments, peptides, which are directly presented by self MHC molecules on the cell surface (Townsend *et al*, 1989). This recognition of "altered self" is carried out by the TCR which can directly bind to pathogen derived peptides in the context of self-MHC molecules. The interaction of TCRs with pathogen peptide MHC complex leads to cell signalling which induces cellular responses such as proliferation, differentiation and cytokine secretion. T cell recognition is dependent upon a whole range of protein-protein interactions that occur between T cell and target cell (van der Merwe and Davis, 2003), culminating in efficient TCR

triggering. In addition to the TCR/MHC interaction itself, the simultaneous binding of coreceptor (CD4 or CD8) to the peptide-MHC molecule is crucial in enhancing the sensitivity of antigen recognition. The coreceptors are able to bind to the non polymorphic region of the MHC molecule, and these interactions are weaker than, and occur after, TCR interaction with the MHC. A number of other interactions, including CD2/CD48, LFA-1/ICAM1 interactions, and costimulatory receptor/ligand interactions are critical in these processes (van der Merwe and Davis, 2003). As a result, TCR triggering occurs, resulting in stimulation of intracellular signalling cascades. Several models have been proposed to explain how TCR triggering occurs, involving either receptor aggregation, receptor conformational change, or kinetic segregation of the TCR-CD3 complex. While these are still debated, a growing body of data support the idea that segregation of TCR/MHC complexes into zones of close contact that exclude the T cell phosphatase CD45 is important in initial triggering, a key point of the kinetic segregation model (Davis and van der Merwe, 2006). Following triggering, active rearrangement of the T cell/target cell interface culminates in formation of the immune synapse (Huppa and Davis, 2003), whereby some of the key cell surface molecules involved are separated into distinct zones at the interface. Although the function of the immune synapse is somewhat unclear, one possibility is that it facilitates targeted secretion of effector molecules to the target cell.

1.3.3 Thymic selection of αβ T cells

TCR gene rearrangement can produce a vast TCR repertoire up to 10^{15} unique αβ TCR sequence combinations, and up to 10^{18} unique γδ TCR sequence combinations, however many of these possible TCR combinations may not be able to recognise self MHC, rendering them useless, or may recognise self-peptide/MHC too strongly, potentially leading to autoimmune

disease. In order to prevent maturation of such ineffective or potentially dangerous specificities, T cells undergo positive and negative selection in the thymus during T cell development (Stritesky *et al*, 2012).

Initially thymocytes are double negative for the expression of the coreceptors CD4 and CD8 and the TCRβ locus undergoes rearrangement. After assembly of a productively rearranged TCRβ chain with the pre-TCRα chain and the CD3 signalling complex, proliferation of the T cell occurs and expression of both CD4 and CD8 coreceptors is upregulated, and at this stage the cell is defined as a double positive T cell. Double positive T cells then undergo positive selection for the ability to recognise self MHC complexes. The majority of TCRs generated cannot recognise self peptide-self MHC and so are not positively selected, eventually dying by neglect (Figure 1.1). The cells which are positively selected mature and express high levels of TCR. At this stage they cease to express one of the coreceptors based on TCR signal strength during positive selection, and become single positive CD4 or CD8 T cells. These single positive T cells express TCRs with a wide range of affinities for MHC molecules. T cells which express potentially high affinity TCRs will undergo negative selection, leading to death by apoptosis (Figure 1.1). This lowers the probability of autoimmune disease and allows survival of T cells that weakly recognise self MHC molecules (Figure 1.1) and have the potential for specificity for foreign antigens in a self-MHC context.

The mature T cell pool is therefore composed of T cells that express TCRs of moderate affinity for self-MHC (Carreño *et al*, 2006). Positive and negative selection of γδ T cells will be discussed later.

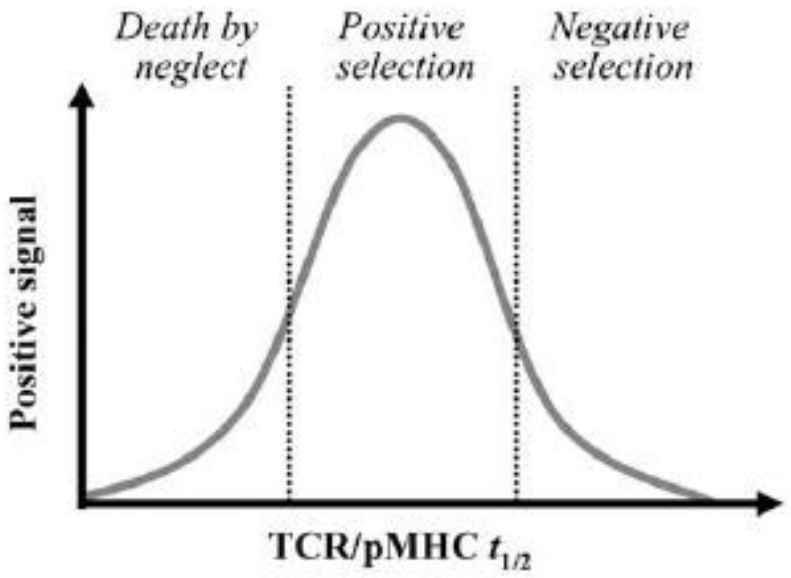


Figure 1.1 Development of T cells is dependent on TCR-pMHC binding kinetics. The model proposed by Carreño *et al*, (2006) suggests that thymocyte positive selection occurs as a result of TCR/pMHC interactions within an intermediate range of half-lives. Short TCR/pMHC interaction half-lives that would prevent completion of TCR signaling due to lack of positive signals and hence thymocytes will die by neglect. However, TCR/pMHC interactions with prolonged half-lives would impede TCR serial engagement impairing TCR signalling, resulting in negative selection and thymocyte death by apoptosis.

1.3.4 Antigen presentation and MHC structure

As mentioned earlier, there are two classes of MHC molecules, MHC class I and MHC class II, which interact with CD8 and CD4 T cells respectively. Importantly, these two classes of molecule are structurally distinct, and the peptides they present are derived from different intracellular locations.

MHC Class I molecules are composed of a heavy or α chain associated noncovalently with a smaller β2-microglobulin (β_2M) domain (Figure 1.2a). There are three MHC class I α chain genes in the human, HLA-A, B and C, and each gene is highly polymorphic, while β_2M is nonpolymorphic. The α chains are transmembrane glycoproteins of approximately 31 kDa and are organised into three extracellular domains (α1, α2 and α3 domains), while the β_2m protein is an Ig-like protein of approximately 12 kDa in size. The association of the heavy chain and β_2M proteins is essential for expression of the MHC on the cell surface. Human MHC class I (specifically HLA-A2) was the first MHC structure to be solved (Bjorkman *et al*, 1987). The structure revealed two pairs of interacting domains, a membrane distal pair consisting of α1 and α2 regions of the heavy chain and membrane proximal pair consisting of α3 region and β_2M (Figure 1.2a). The α1 and α2 domains interact to produce a peptide binding platform which contains two α-helices which rest on a β-sheet, and form a groove in which the peptide is sited. This groove can accommodate a peptide up to ten amino acids in length. The α3 domain and β_2M are Ig-like domains of two anti-parallel β-sheets. In contrast to the α1 and α2 domains, which are highly polymorphic, the α3 domain is relatively conserved across MHC class I molecules and is a key interaction site for the CD8 T cell coreceptor (Gao *et al*, 1997).

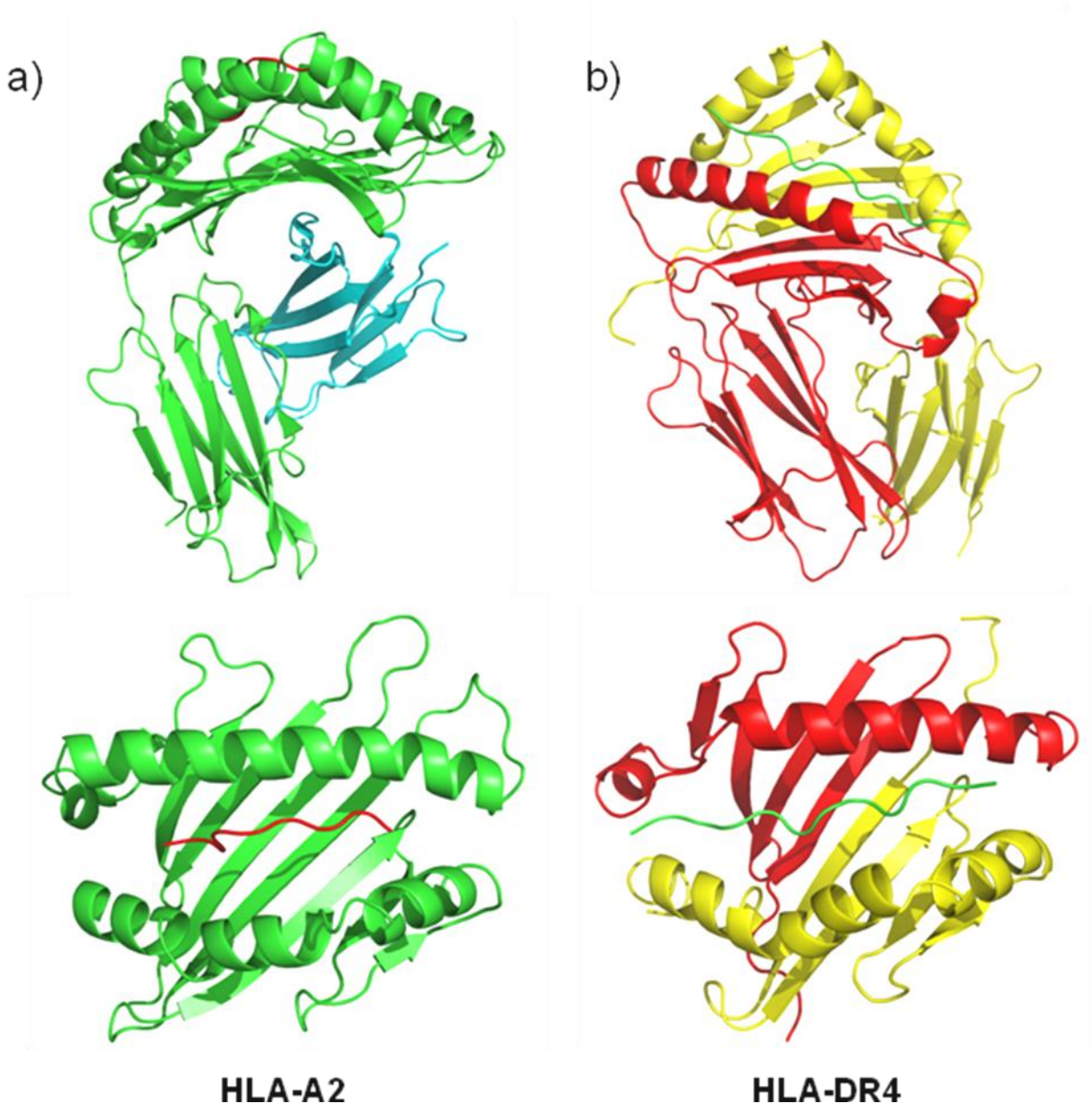


Figure 1.2 Overview of the crystal structures of MHC-I (HLA-A2) and MHC-II (HLA-DR4). a) Structure of HLA-A2 in complex with peptide derived from Influenza A matrix protein M1 (PDB code 1HHI; Madden *et al*, 1993). The α1 and α2 domains (green) combine to form the antigen binding cleft (bottom) allowing peptide (red) presentation to $CD8^+$ T cells. The α3 domain (green) non-covalently associates with β_2M (cyan). b) Structure of HLA-DR4 in complex with peptide derived from human collagen II (PDB code 2SEB; Dessen *et al*, 1997). MHC-II molecules are comprised of a polypeptide alpha (red) and beta (yellow) chain of approximately equal length. The α1 and β1 domains combine to form the antigen binding groove allowing peptide (green) presentation to $CD4^+$ T cells. The groove of MHC-II is open at one end relative to MHC-I allowing accommodation of longer peptides (13-25).

The assembly of MHC class I molecules begins in the endoplasmic reticulum (ER), where newly synthesized MHC class I heavy chain and β_2M assemble with each other in association with a large multiprotein complex termed the Peptide Loading Complex (PLC) (Pamer and Cresswell, 1998; Cresswell *et al*, 1999). The peptides which are loaded onto the MHC class I complex are generated in the cytosol from cytosolic proteins that have been degraded by a multi-subunit protease called the proteasome. The peptides generated by the proteasome are then transported into the ER by the transport associated with antigen processing (TAP) complex. Peptides that associate stably with the MHC class I- β_2M dimers are ultimately transported to the cell surface (Pamer and Cresswell, 1998).

MHC class II are structurally similar to MHC class I molecules, but consist of two polypeptide chains of approximately equal size, which are associated by noncovalent interactions (Figure 1.2b). The two chains consist of two domains each: α chain has α1 and α2 domains and the β chain has β1 and β2 domains. The α1 and β1 domains fold together to form the peptide binding groove, analogous in structure to that of MHC class I, however the ends of the peptide binding groove are not constrained, and as a result, MHC class II proteins are able to bind longer peptides (Figure 1.2b). Nascent MHC class II molecules are initially assembled in the ER, as for MHC class I molecules (Watts, 1997). However, unlike MHC class I, the binding groove of MHC class II molecules is occupied at first by association with a portion of the Class II MHC invariant chain (Ii). This association prevents peptides in the ER from binding to the MHC class II molecule, and helps traffic the MHC class II molecule out of the ER towards the endocytic pathway, where it ultimately intersects with endosomal vesicles carrying antigen originating from outside the cell. In the low pH of these vesicles, the Ii chain is degraded by proteases, leaving the Class II-associated invariant chain (CLIP)

peptide bound to the MHC class II peptide binding groove. After fusion with endocytic vesicles, the CLIP peptide is displaced from the MHC class II groove by peptides derived from endocytosed antigens that have been similarly proteolytically degraded. Exchange of the CLIP peptide for antigenic peptide is catalysed by the class II-like molecule HLA-DM protein, after which mature MHC class II complexes are transported to the cell surface for presentation to T cells (Watts, 1997).

1.3.5 Structural basis of TCR/MHC recognition

Structural studies of TCRs and TCR/MHC complexes have shed light on how T cell recognition takes place at a molecular level (Rudolph *et al*, 2006). TCR structures have confirmed structural homology to antibody Fab fragments, with antigen binding site formed by the CDR loops protruding from the membrane distal ends of the Vα and Vβ domains. Generally, TCR/peptide-MHC structures have revealed a roughly diagonal orientation of the TCR relative to the long axes of the α1/α2 helices, with the Vα domain positioned over the N-terminal portion of the peptide (and the α2 helix), and the Vβ domain positioned over the C-terminal region of the peptide (and the α1 helix). In this orientation, the CDR3 loops of both chains are positioned over the centre of the antigen binding groove and dominate contacts to antigenic peptide, whereas contacts to α1 and α2 helices are dominated by the CDR1 and CDR2 loops of Vβ and Vα respectively. Importantly no gross conformational changes in either TCR or MHC have been noted upon binding, arguing against models of TCR triggering based on conformational change. Furthermore, crystallographic studies have shown that TCRs can accommodate altered peptide ligands of diverse stimulatory capacity with only minor conformational adjustments (Ding *et al*, 1999). Similarly, crystal structures have provided no evidence for TCR triggering models based on ligation-dependent dimerisation. Instead, both

binding studies and structural studies have highlighted relatively minor conformational alterations at the TCR/MHC protein-protein interface, particularly involving conformational changes in CDR loops upon binding (Willcox *et al*, 1999; Boniface *et al*, 1999; Rudolph *et al*, 2006). Collectively, these data are arguably most consistent with models of triggering based on kinetic segregation (Davis and van der Merwe, 2006).

1.4 Development of the adaptive immune response

A wide range of studies provide support for a model of conventional adaptive responses based on initial activation of myeloid cells, followed by activation and recruitment of antigen specific lymphocytes, which takes place after a considerable delay (Hayday, 2009). After initial PRR stimulation, myeloid cells help to activate innate responses, but also upregulate antigen presentation molecules and migrate to local lymphoid tissues, where they present antigens to T cells. After initial priming, antigen-specific T cells have to proliferate, differentiate into effector T cells, and migrate to the affected site in order for their potent adaptive effector responses to be carried out. In addition, a portion of the effector cells are retained in a long-lived memory pool. In this way, subsequent exposure to the same antigen/pathogen enables rapid differentiation to produce a wave of effector cells that mount a quick secondary response (Hayday, 2009). While this system of adaptive immune initiation is clearly very powerful and can be sufficient to eliminate some pathogen infections, the time taken for initiation of antigen-specific lymphocyte responses suggests that in some cases, other arms of the immune response may be required to provide fast responses to microbial and non-microbial challenges. Increasing evidence suggests unconventional lymphocytes may fulfil this role (Hayday, 2000; Hayday, 2009).

1.5 Unconventional lymphocytes

Additional subsets of T cells, classed as unconventional T cells, do not recognise self MHC molecules and peptides. These cells can express an αβ TCR and may recognise MHC-like molecules or unrelated antigens. A major class of unconventional αβ T cells recognises CD1 molecules presenting antigenic lipids (Brigl and Brenner, 2004), while another class is specific for the MR1 molecule (Tilloy *et al*, 1999; Treiner *et al*, 2003) CD1 molecules are homologous to MHC class I molecules, but have evolved specialised hydrophobic antigen binding grooves, enabling presentation of lipid antigens to T cells (Figure 1.3a). In humans five CD1 genes are present, with CD1d (as in mouse) being recognised by Natural Killer T cells, which express a highly restricted αβ TCR and have potent immunoregulatory capacity. In contrast, the group 1 CD1 molecules (CD1a, b, c) present in humans are recognised by αβ T cells with diverse TCRs, and are implicated in responses to mycobacterial lipids (Willcox *et al*, 2007). In addition, muscosal associated invariant T cells (MAIT) have a highly restricted TCR usage, and recognise the class I MHC-like MR1 molecule, which has recently been shown to present a folic acid metabolite to αβ T cells (Kjer-Nielsen *et al*, 2012) (Figure 1.3b). In addition, γδ T cells undoubtedly represent one major class of unconventional T cell (Hayday, 2000), and these are discussed in greater detail in subsequent sections.

Collectively, unconventional T cells, together with Natural Killer cells that also belong to the lymphocyte lineage, demonstrate the potential to respond rapidly to complement the conventional adaptive response, by responding rapidly to signs of either microbial challenge, or non-microbial stress such as cellular transformation. Collectively, these processes have been termed “lymphoid stress surveillance” (Hayday, 2009) and will be considered in more detail below in the context of γδ T cells.

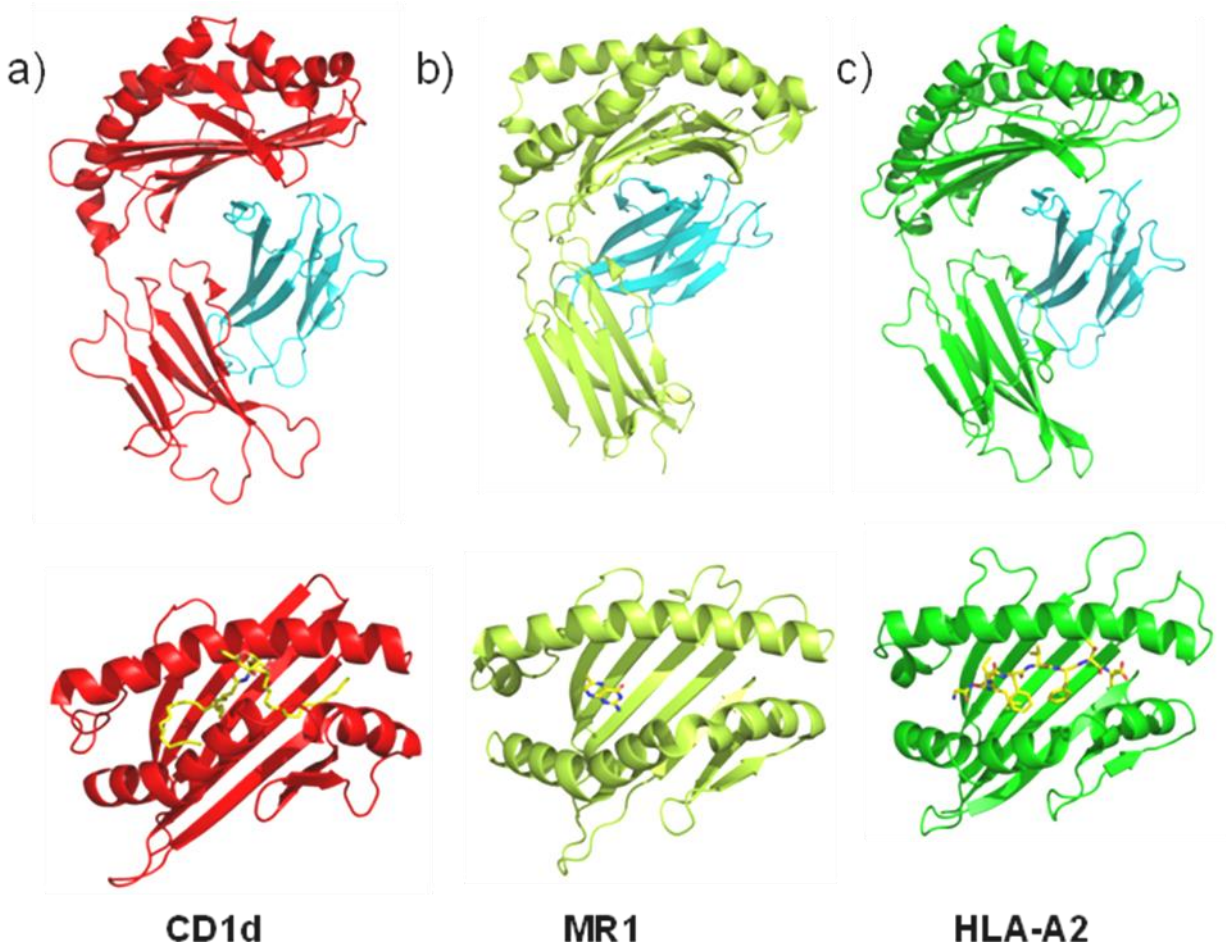


Figure 1.3 Established ligands for unconventional αβ T cells. a) Crystal structure of CD1d bound to α-galactosylceramide (PDB code 1ZT4; Koch *et al*, 2005). The α1 and α2 domains (red) combine to form the antigen binding cleft (bottom) in which the lipid (ball and stick format) is presented to invariant natural killer T cells. The α3 domain (red) non-covalently associates with β_2M (cyan). b) Crystal structure of MHC-I related molecule MR1 in complex with 6-formyl pterin (PDB code 4GUP; Kjer-Nielsen *et al*, 2012). The α1 and α2 domains combine to form the antigen binding cleft in which the microbial vitamin B9 metabolite (ball and stick format) is presented to mucosal-associated invariant T (MAIT) cells. The α3 domain (lime) non-covalently associates with β_2M (cyan). c) For comparison the crystal structure of HLA-A2 in complex with peptide derived from Influenza A matrix protein M1 (PDB code 1HHI; Madden *et al*, 1993) is shown. The α1 and α2 domains (green) combine to form the antigen binding cleft in which the peptide (ball and stick) is presented to $CD8^+$ T cells. The α3 domain (green) non-covalently associates with β_2M (cyan).

1.6 An introduction to γδ T cells

One important feature of the adaptive immune response, as outlined above, is the considerable time required from initial antigen encounter to generation of antigen-specific lymphocyte effector populations. This emphasises the requirement for a complementary arm of the immune system, to mount rapid and initial responses. In addition to innate immune responses, growing evidence suggests unconventional lymphoid subsets make key contributions at these early stages, acting to limit infection, initiate and regulate downstream responses, and preserve tissue integrity. Collectively, these early lymphoid-mediated responses have been referred to as "lymphoid stress surveillance" (Hayday, 2009). In addition to Natural Killer cells, γδ T cells are emerging as arguably the archetypal lymphocyte subset involved in this process.

1.6.1 The discovery of γδ T cells

The existence of γδ T cells as a third lineage of lymphocytes alongside conventional T cells and B cells was completely unanticipated, and stemmed from a serendipitous discovery. In their attempt to identify the α chain of the TCR, Saito *et al* instead cloned a TCR γ cDNA (Saito *et al*, 1984). A key subsequent step was to understand how the γ chain was encoded, and studies soon afterwards outlined the organisation of the γ locus on chromosome 13, which is arranged in four clusters, each containing one or more V segments, and both a single J segment and single C gene (Hayday *et al*, 1985; Vernooij *et al*, 1993). A second key question was whether the γ chain paired with another TCR chain to form a heterodimer, as for the αβ TCR. A number of biochemical studies, both in mouse and human systems, provided evidence that a fourth chain did indeed heterodimerise with γ (Brenner *et al*, 1986; Pardoll *et al*, 1987; Nakanishi *et al*, 1987). Subsequent work revealed that the murine Vδ locus was

embedded within the α chain locus, between the Vα and Jα segments, on chromosome 14 (Chien *et al*, 1987). Interestingly, there are relatively few Vδ segments compared to Vα gene segments, however Vα genes can be utilised during γδ TCR recombination, increasing the potential diversity of the Vδ chain (Elliott *et al,* 1988).

1.6.2 γδ TCR recombination and diversity

A striking feature of the γδ TCR is the paucity of Vδ and Vγ gene segments compared to Vα and Vβ segments (Table 1.1). Furthermore, several γ and δ chains preferentially pair with each other. Although both of these features serve to limit combinatorial diversity for the γδ TCR and might suggest overall genetic diversity within the γδ TCR is also low, this is not the case. In fact, two features of TCR δ recombination have the effect of increasing diversity substantially relative to the αβ lineage (Davis and Bjorkman, 1988) (Table 1.1). Firstly, although diversity (D) regions are not utilised in recombination of the Vγ chain (as is also the case for the Vα chain), each δ chain can utilise more than one Dδ region, whereas TCR β chains use only one D region. Secondly, these Dδ segments can be read in any frame. As a result of these features, the potential diversity of the γδ TCR exceeds that of either the BCR repertoire or the αβ TCR.

Molecular features of the TCR γ and δ chain also distinguish these from α and β TCR chains to an extent. One notable feature is that although human and mouse Vh, Vk, Vα, Vβ and Vδ gene segments can be separated into distinct families that each share relatively high homology at the nucleotide level (>75%), Vγ gene segments are by contrast genetically diverse. Although Vγ1 (human) and Vγ7 (mouse) exhibit >40% homology, other human and mouse Vγ genes are relatively distinct. In addition, it has been noted that although Vδ gene segments

are homologous to those of Vα segments, the "classical" Vδ segments used exclusively by Vδ chains (Vδ1, Vδ2, Vδ3 in humans, and Vδ1, Vδ4, Vδ5, and Vδ6 in mice), as opposed to those shared by Vα chains, bear key sequence features distinct from Vα genes (Arden *et al*, 1995a; Arden *et al*, 1995b; Clark *et al*, 1995). These include a characteristic increased length (by two amino acids) of CDR1δ and 2δ regions relative to CDR1α and 2α regions. In addition, detailed comparisons have revealed certain amino acids that are highly conserved in Vα genes such as glutamate 14 (hypothesised to be involved in contacting the Cα domain) and glutamine 37 (hypothesised to form contacts to the β chain) are not conserved in Vδ.

	Immunoglobulin		TCR α : β		TCR γ : δ	
	H	**K**	**α**	**β**	**γ**	**δ**
Variable segments	250-1000	250	100	25	7	10
Diversity segments	10	0	0	2	0	2
Ds read in all frames	Rare	-	-	Often	-	Often
N region addition	V-D, D-J	None	V-J	V-D, D-J	V-J	V-D_1, D_1-D_2, D_1-J
Joining segments	4	4	50	12	2	2
V region combinations	62,500 – 250,000		2,500		70	
J combinations	~10^{11}		~10^{15}		~10^{18}	

Table 1.1 Sequence diversity in immunoglobulin genes and αβ/γδ TCRs (Davis and Bjorkman, 1988.

1.6.3. Essential features of γδ T cell localisation and development

γδ T cells can be separated into two subtypes. The first is systemic γδ T cells, which comprise 1-10% of circulating T cells in both mice and humans. Secondly, tissue-associated γδ T cells are enriched within epithelia, and are termed "intraepithelial γδ lymphocytes" or IELs. Their prevalence at epithelial surfaces in particular suggests γδ T cells may act as a first line of

defense against infection (Hayday, 2000). In the mouse, IELs demonstrate restricted V-region usage and limited junctional diversity, with particular combinations characteristic of distinct anatomical sites. As examples, uterine γδ's utilise an invariant Vγ6Vδ1 TCR, whereas intestinal γδ's predominantly use Vγ7, and skin DETC γδ's use an invariant Vγ5Vδ1 combination (Figure 1.4a). In contrast, systemic γδ T cells are more diverse in chain usage and have higher junctional diversity. Although human γδ T cell subsets are less well characterised, the predominant peripheral blood γδ subset utilises Vγ9Vδ2. In contrast, tissue-associated γδ T cells are predominantly Vδ2-negative, with use of Vδ1 highlighted in the intestine, and generally diverse Vγ chain usage (Figure 1.4b).

In both mice and humans, γδ T cells are the first T cells to develop, suggesting important roles in neonatal immunity. In mice, their development occurs in waves, with DETC appearing first, followed by uterine and then other γδ T cell populations (Figure 1.5). The lack of junctional diversity in DETC and uterine populations can be explained in part by the fact their recombination occurs before terminal deoxynucleotide transferase (TdT) is expressed, however some studies also suggest canonical rearrangements may either be enriched by positive selection, or may occur at high frequencies due to homologous recombination directed by short nucleotide sequences flanking the V gene segments. For the DETC TCR, evidence for positive selection derives from disruption of the Vγ5 gene used by DETCs (Mallick-Wood *et al*, 1998). In normal mice, the canonical DETC TCR Vγ5/Vδ1 is recognised by the 17D1 antibody, which does not recognise either Vγ5 or Vδ1 paired with other chains. In the skin of $V\gamma5^{-/-}$ mice, DETCs were present but they bore TCRs comprising a replacement Vγ1 chain paired with the canonical Vδ1 chain. Interestingly, in $V\gamma5^{-/-}$ mice a significant proportion of the replacement DETC were reactive against the 17D1 antibody,

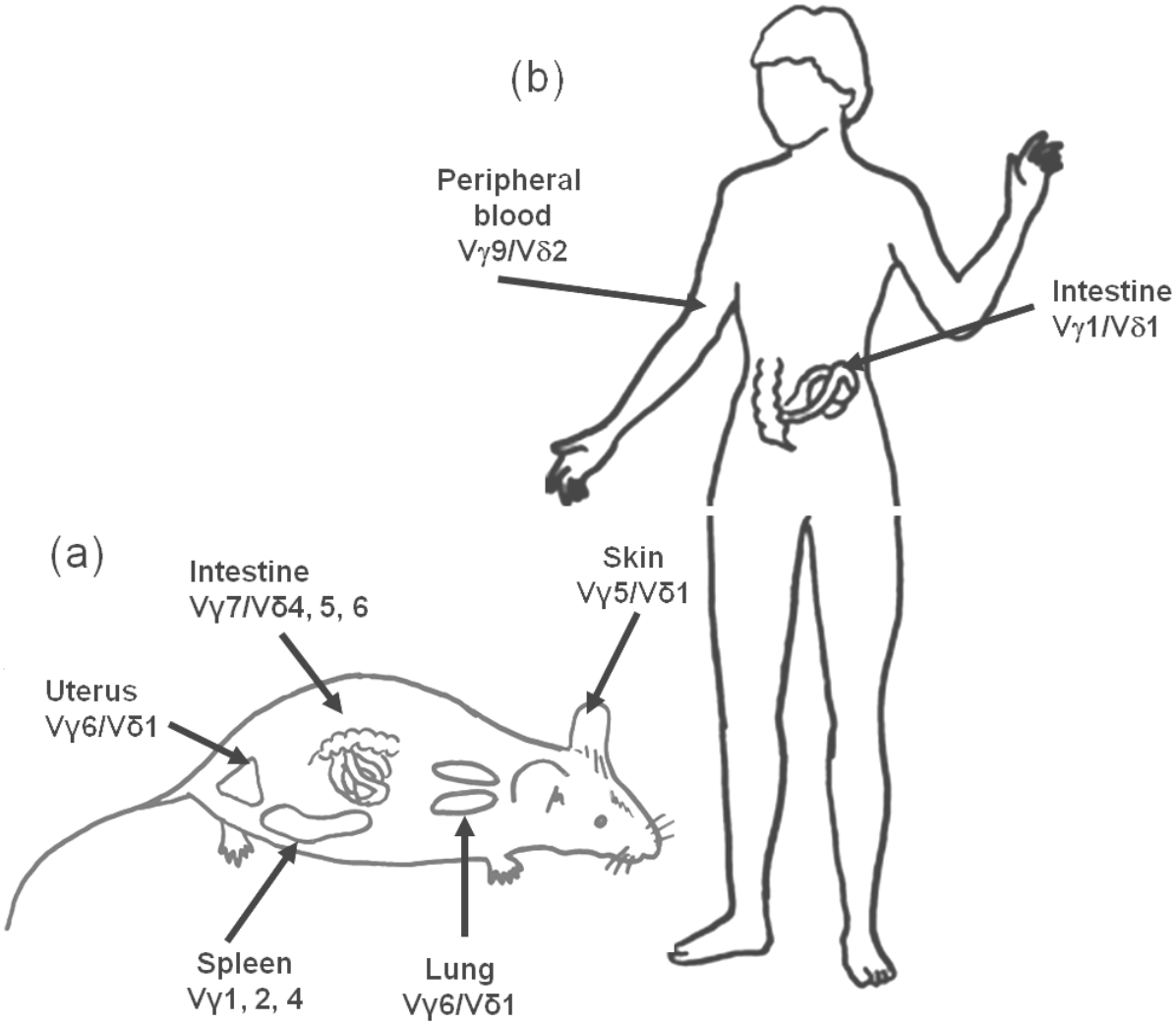


Figure 1.4 γδ gene usage in mouse (a) and human (b). In each case γδ T cells show limited V region usage at particular anatomical sites.

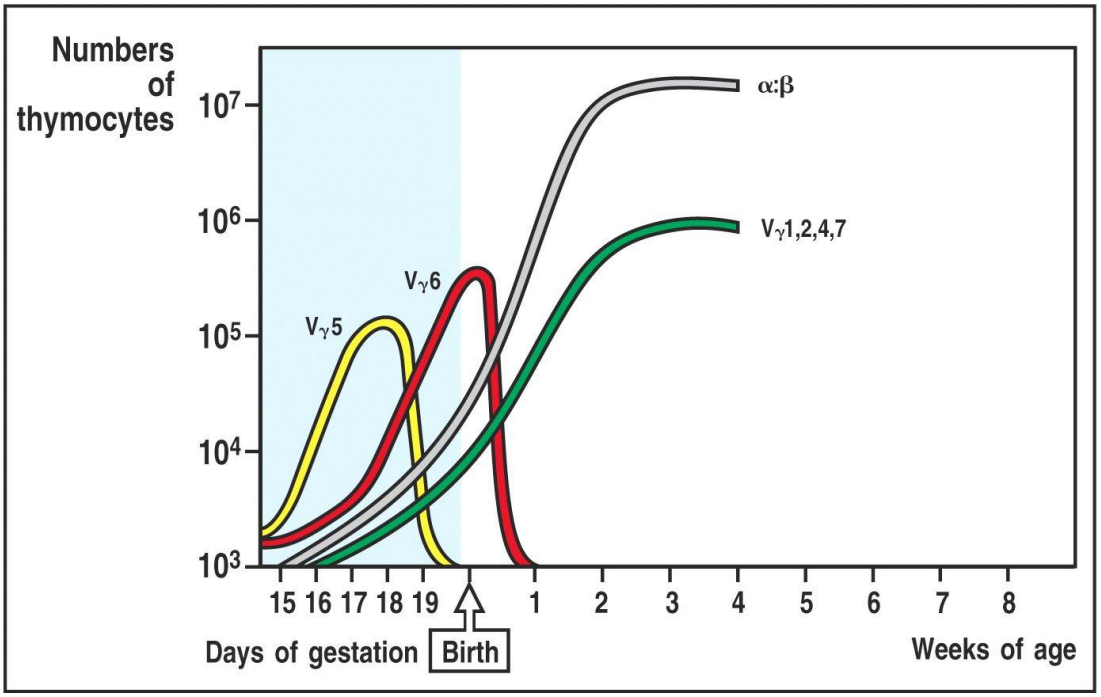


Figure 1.5 Vγ gene segment usage is developmentally regulated in the murine thymus. The first wave of thymocytes produced in the embryonic thymus rearranges Vδ5 and migrates to the skin, while the next wave rearranges Vδ6 and homes to the lung and uterus. αβ T cells and peripheral and intestinal γδ cells, which are more diverse, undergo rearrangement and selection following birth (Janeway *et al*, 2005).

indicating that many Vγ1Vδ1 TCRs could adopt a conformation similar to that of the canonical Vγ5Vδ1 DETC TCR, and suggestive of thymic positive selection events based on recognition of a ligand that may be also present in the skin.

An important difference between γδ and αβ T cell development that is highly relevant to the points above, relates to the pre-TCR. Whereas in αβ T cell development, functional expression of a pre-TCR is required to rescue cells from apoptosis at the DN stage immediately prior to the DP $CD4^+CD8^+$ stage, there is no evidence of a functional pre-TCR checkpoint for γδ T cells, the majority of which mature as DN cells. The assumption is that like the pre-TCR (Yamasaki *et al*, 2006), the majority of γδ TCRs can signal independently of thymic ligands sufficiently to effect maturation. Recent evidence in mice supports this idea, indicating that most Vδ chains can substitute for the pTα chain. Importantly, the one exception to this observation was the Vδ1 DETC TCR chain, consistent with the idea that interaction of the canonical DETC TCR with a thymic ligand is an essential positive selection step (Jensen *et al*, 2008).

Finally, whereas negative selection of αβ T cells strongly reactive against self-MHC is well established, this is not generally thought to occur with γδ T cells. The most convincing evidence for this comes from studies on mouse γδ T cells reactive against the non-classical class I MHC molecules T10 and T22. Although early studies involving mice transgenic for the G8 TCR that recognises T10/T22 favoured both positive and negative selection (Dent *et al*, 1990; Bonneville *et al*, 1990; Perreira *et al*, 1992; Wells *et al*, 1991), the use of T10/T22-specific tetramers detected comparable frequencies of T10/T22-specific γδ T cells was also found in $\beta_2M^{-/-}$ mice, excluding negative (and positive) selection playing a major role

(Crowley *et al*, 2000). Although this is somewhat surprising given the relatively high affinity of some γδ TCR/T10-T22 interactions, the observation that there is an unexpectedly high threshold for activation in spite of this may provide an explanation. Instead, in the T10/T22 system, the presence or absence of TCR ligand in the thymus is suggested to affect T_H1/T_H17 phenotype bias, with thymic selection in the presence of ligand inducing a T_H1 phenotype, and vice-versa (Jensen *et al*, 2008). Moreover, in humans, results from next generation sequencing analyses comparing frequencies of in-frame and out-of-frame Vγ genes have been used to suggest that only low levels of negative selection may occur (Sherwood *et al*, 2011). In summary, there is strong evidence against negative selection playing a dominant role in shaping the γδ T cell repertoire, unlike for αβ T cells. The implications of this are highly significant. Firstly, they suggest that mature γδ T cells may have the potential to recognise self antigens, even if these are expressed in the thymus. Secondly, they raise the key question of how γδ T cell tolerance to self-antigens is maintained in the absence of negative selection.

1.6.4. Key γδ T cell subsets

Much of the research on γδ T cells has been focussed on key γδ T cell subsets. A brief summary of some key γδ subsets is provided below.

Vγ9Vδ2 T cells are restricted to humans and primates and are the predominant population in human peripheral blood. Known to expand rapidly in a range of infections, including mycobacterial infection, salmonella, listeria, toxoplamosis, and malaria, Vγ9Vδ2 have been described to be activated by small non-peptidic phosphorylated compounds derived from bacterial-specific isoprenoid biosynthetic pathways, commonly referred to as phosphoantigens (Pfeffer *et al*, 1990; Constant *et al*, 1994; Tanaka *et al*, 1994). The most potent of these,

HMB-PP, could potentially be considered a PAMP, as it is an essential metabolite for most pathogenic bacteria such as *Mycobaterium tuberculosis*, and is produced via the non-mevalonate pathway of isoprenoid synthesis absent in higher eukaryotes. Therefore, Vγ9Vδ2 T cells may play an important role in sensing infection by HMB-PP-producing pathogens. Importantly, host-produced phosphoantigens such as isopentenyl pyrophosphate (IPP) are able to stimulate Vγ9Vδ2 cells, these are 10^4-10^7-fold less potent than HMB-PP, which has a bioactivity of 0.1nM (Eberl *et al*, 2003). Conceivably, molecules such as IPP may promote recognition of stressed or cancerous cells, supported by *in vitro* reactivity of Vγ9Vδ2 T cells to transformed tumour cell lines. In addition, studies by Moser *et al* highlight the potential for Vγ9Vδ2 cells to act as antigen presenting cells to αβ T cells, following phosphoantigen stimulation (Moser *et al*, 2006). The direct anti-tumour reactivity of Vγ9Vδ2 cells is potentially manipulable in the therapeutic setting, as aminobisphosphonates (e.g. zoledronate and pamidronate) used clinically to decrease bone loss (e.g. in myeloma) stimulate Vγ9Vδ2 T cells by inhibiting enzymes of the classical mevalonate pathway, leading to IPP accumulation (Hewitt *et al*, 2005). A number of clinical trials have used aminobisphonates to elicit γδ T cell-mediated anti-tumour responses, with some evidence of efficacy (Wilhelm *et al*, 2003 Dieli *et al*, 2003). Both for HMB-PP and IPP, the mode of recognition is unclear, however a species-specific antigen-presentation molecule is hypothesised (Hewitt *et al*, 2005).

Human Vδ2-negative γδ T cells are more poorly characterised than the $V\delta2^+$ peripheral blood subset. Although they are predominantly tissue-associated, Vδ2-negative cells are present in the peripheral blood at relatively low frequencies (typically <2% of circulating T cells), and expand upon certain infections, including HIV, Lyme disease, and CMV. The response to CMV is arguably best characterised, and involves diverse TCR chain usage, that varies from

patient to patient and involves clonotypic expansions (Déchanet *et al*, 1999). In addition, expanded populations of CMV-specific Vδ2-negative cells bearing an effector memory phenotype persist post-infection, highlighting similarities to adaptive responses (Pitard *et al*, 2008). Within the Vδ2-negative subset, $V\delta1^+$ cells have been most extensively studied, and their reactivity to epithelial tumour cells emphasised (Groh *et al*, 1998; Groh *et al*, 1999). Moreover, CMV-reactive Vδ2-negative cells can exhibit a dual reactivity against tumour cells, suggesting reactivity towards self stress antigens (Halary *et al*, 2005). In both cases, the effector functions of the Vδ2-negative cells include cytokine production and cytotoxicity. The ligands recognised by Vδ2-negative are unknown, although they are clearly distinct from the phosphoantigens that stimulate the Vγ9Vδ2 subset. One suggested TCR ligand is the class I MHC-like protein MICA (Li *et al*, 1999; Wu *et al*, 2002) (Figure 1.6), however this remains controversial, particularly as these cells universally express NKG2D, a high affinity ligand for MICA. The Vδ2-negative population is discussed more extensively below.

A key population in mice is the skin-resident intra-epithelial lymphocyte population known as dendritic epidermal T cells, or DETC (Hayday, 2000). These cells express a highly invariant Vγ5Vδ1 TCR, and in the steady state, adopt a dendritic morphology, contacting numerous keratinocytes in the epidermal layer. DETC are able to respond to keratinocytes *in vitro*, and this is augmented by exposing keratinocytes to stress stimuli such as heat shock. The responses induced include production of IL-2, cytotoxicity, production of keratinocyte growth factor, and secretion of chemokines (e.g. lymphotactin) that can recruit $CD8^+$ T cells. Consistent with this, studies in mice support a role for DETC both in promoting epithelial tissue repair, and also in cutaneous tumour immunosurveillance. DETC are discussed in greater detail below.

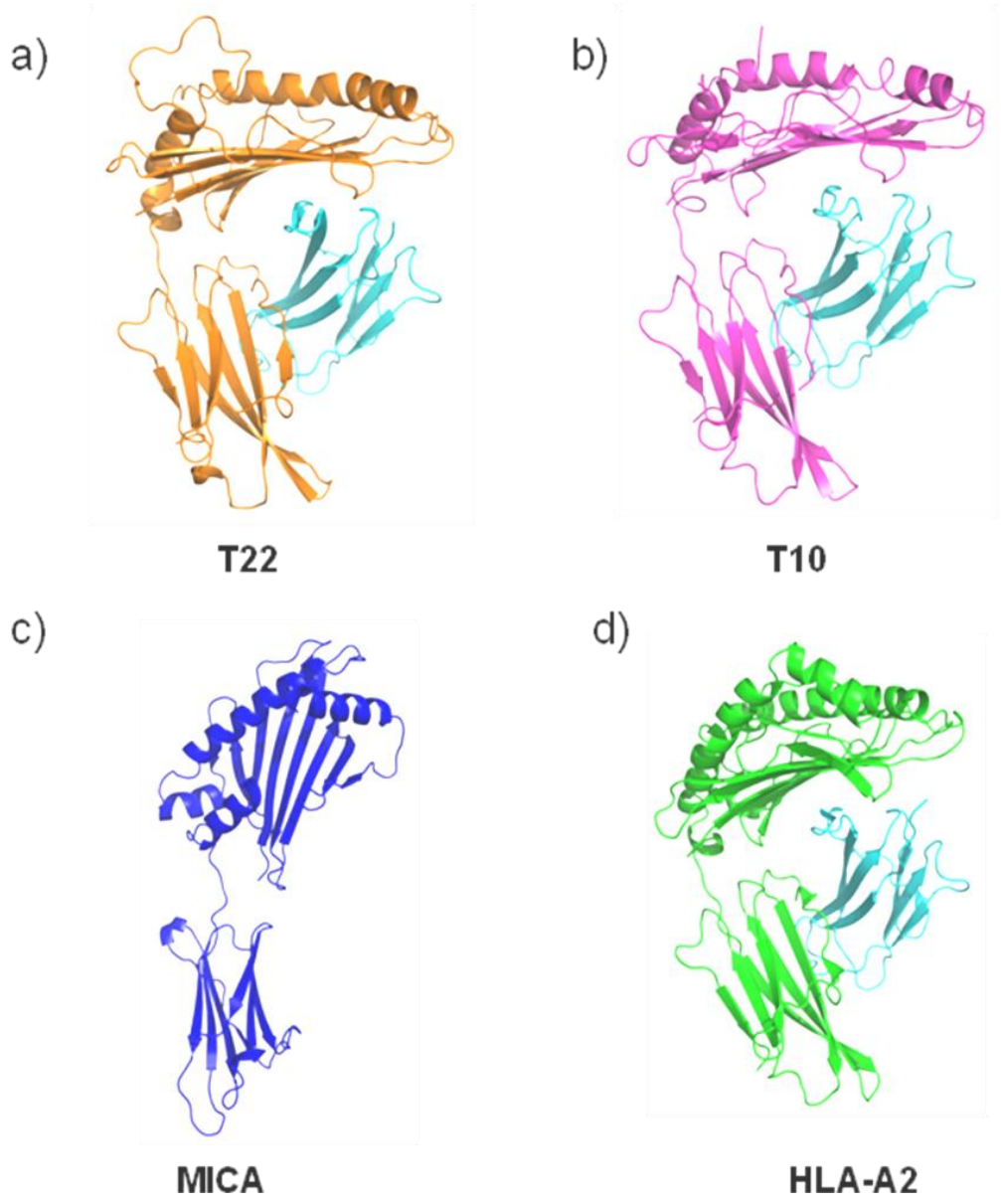


Figure 1.6 Established ligands for unconventional γδ T cells. a) Crystal structure of murine T22 (PDB code 1C16; Wingren *et al*, 2000). The α1 and α2 domains (orange) combine to form a MHC-I like antigen binding groove. The α3 domain (orange) non-covalently associates with β_2M (cyan). b) Crystal structure of murine T10 (PDB code 1R3H; Rudolph *et al*, 2004). The α1, α2 domains (pink) combine to form a MHC-I like antigen binding groove. The α3 domain non-covalently associates with β_2M (cyan). Both T10 and T22 are recognised by γδ T cells in the absence of endogenous/exogenous ligands. c) Crystal structure of MHC class I polypeptide-related sequence A (MICA) (PDB code 1B3J; Li *et al*, 1999). The α1, α2 and α3 domains are highlighted in blue. MICA serves as stress inducible ligand that is recognized by a subset of γδ T-cells (expressing Vδ1) independent of β_2M and bound peptides. d) For comparison the crystal structure of HLA-A2 in complex with peptide derived from Influenza A matrix protein M1 (PDB code 1HHI; Madden *et al*, 1993) is shown. The heavy chain (α1, α2 and α3 domains) is shown in green β_2M in cyan).

γδ T cell reactivity to the mouse class IB (ie non-classical) MHC molecules T10 and T22 was first defined in two alloreactive T cell hybridomas, G8 and KN6. T10 and T22 are highly related (sharing 94% identity), and associate with β2-microglobulin (Figure 1.6), although neither binds an antigen (Wingren *et al*, 2000). As is the case for MICA, T10/T22 expression is increased on activated cells, and a threshold of T10/T22 is required for G8/KN6 activation. While the physiological significance of T10/T22 recognition is unclear, the T10/T22 system has proven a valuable system in which to examine in particular γδ T cell selection (described above), and γδ TCR ligand recognition (discussed in more detail below).

1.7 Functional roles of γδ T cells

1.7.1 Evolutionary conservation and the lymphoid stress surveillance response

The γδ T cell subset appears to have co-evolved alongside αβ T cells and B cells, since early in vertebrate evolution, which is around 400 million years ago, around the time when jawed vertebrates emerged. This suggests a critical and fundamental role for γδ T cells in vertebrate immunity, and one that is likely to be complementary to conventional αβ and B cell responses (Hayday, 2000). Consistent with such a fundamental role, to date no humans have been observed that specifically lack γδ T cells. However, the key reasons underlying this evolutionary conservation have until recently, remained unclear. This is illustrated by the observation that many γδ T cell knockout models are unremarkable in terms of the phenotypes observed. Despite this, their enrichment at epithelia surfaces fits well with a potential role for γδ T cells in the "first line of defense", a role that would complement adaptive immunity. Consistent with this, a wealth of evidence (summarised below, with important contributions from animal model studies), now suggests that γδ T cells play key and in some cases essential roles in immune responses to both microbial and non-microbial stress, and illustrate many of

the key features of lymphoid stress surveillance. To understand these processes it is useful to consider the effector responses γδ T cells can carry out.

1.7.2 Effector responses of γδ T cells

γδ T cells exhibit a diverse range of effector functions, as for αβ T cells. These include cytotoxicity, consistent with their expression of cytolytic effector molecules such as perforin, granzymes, and Fas/FasL (Hayday, 2000). Cytokine production is another key function, with IFNγ and TNFα being prevalent. Previous studies have characterised γδ T cells as either T_H1 (as defined by IL-2 and IFNγ production) or T_H2 (as defined by IL-4, IL-5, IL-6. IL-10 production) in phenotype, with the majority conforming to the T_H1 phenotype, and the T_H2 γδ T cell subset correlating with rare expression of CD4 (Wen *et al*, 1998; Wen *et al*, 1997). Importantly, comparisons of the transcriptome of mouse intestinal γδ IELs versus αβ IELs suggested a T_H1-skewed "activated but resting" phenotype, comprising naive expression of a range of cytolytic markers, chemokines, chemokine receptors, and adhesion molecules (Shires *et al*, 2001). In contrast to αβ T cells, these γδ cells expressed various proteins that inhibit proliferation. More recently, IL-17-producing γδ T cells have been described (Roark *et al*, 2008). In some model systems, such as bacterial infection, γδ T cells can provide an innate, rapid, substantial and potentially protective source of IL-17 prior to generation of classical adaptive αβ T_H17 cells (Lockhart *et al*, 2006, Umemura *et al*, 2007, Martin *et al*, 2009; Sutton *et al*, 2009). In addition, IL-17-production by γδ T cells has been highlighted as playing a critical role in chemotherapy-induced anti-cancer immune responses (Ma *et al*, 2011). Finally, a third important effector function of some γδ T cells is production of key growth factors, exemplified by skin-resident Vγ5Vδ1 DETC, which make keratinocyte growth factor upon activation (Boismenu and Havran, 1994).

1.7.3 Immunoprotection from infection

Observations that expansions of γδ T cells occur in a diverse range of pathogen infections (as detailed above), combined with *in vitro* analyses of effector function, provide strong but circumstantial evidence for their importance in immunoprotection against infectious challenge (Hayday, 2000). Complementary to these approaches are γδ T cell knockout studies in mice, which have directly addressed the importance of γδ T cells in various infectious models. These studies have highlighted some examples (detailed below) where γδ T cells provide essential contributions to anti-pathogen immunity. In other cases their effector responses are either insignificant, or redundant (as in Eimeria infection of the gut), whereby deficiency in the γδ T cell subset in adults can be compensated for by αβ T cells (Hayday *et al*, 2000; Ramsburg *et al*, 2003). Therefore collectively, published data in this area highlight that the relative contribution of γδ T cells to immunoprotection against infection is dependent on the exact pathogen involved.

A key example of an essential role for γδ T cells is infection of the lung with *Nocardia asteroids*, an intracellular $Gram^+$ bacteria (King *et al*, 1999). In wild type animals, which all survived infectious challenge, bacteria were cleared from infected airway epithelial cells by an inflammatory response driven by neutrophils. However, all γδ negative mice died, due to a failure to develop inflammatory infiltrates. This highlights the potential role of γδ T cells in protection of an epithelial barrier, and their importance in orchestrating downstream inflammatory responses (King *et al*, 1999). A second example is lung infection with the Gram negative bacteria *Klebsiella pneumonia* (Moore *et al*, 2000). In this case, γδ negative mice have considerably higher mortality compared to wild type mice. Although, initial bacterial clearance in the lung is similar, a disseminated infection results, correlating with lower IFNγ

and TNFα levels, and resulting in liver damage. Consequently, γδ T cells are critical in this model for early production of IFNγ and TNFα; in contrast, αβ negative mice are no more susceptible to infection than wild type mice. A third example is *Listeria monocytogenes*, where knockout studies support a role for γδ T cells in primary infection, as $\gamma\delta^{-/-}$ mice have massively enhanced bacterial titers. Their key role in this scenario appears to be early production of IFNγ, which, by stimulating monocytes to produce TNFα, acts to enhance the NK IFNγ response. This example illustrates the potential of γδ T cells to orchestrate downstream innate immune responses. Collectively, these examples illustrate the principles that γδ T cells can play critical roles in host immune defence against pathogens, and these functions are most critical early in infection, where their rapid responses serve to protect epithelial surfaces, and initiate and shape downstream innate and adaptive responses.

In contrast to the models described above, other systems highlight γδ T cell contributions that, while potentially important, are compensated for by the presence of αβ subsets, and therefore in the αβ replete setting, redundant. One example is HSV-1 infection, where both αβ negative and γδ negative animals are mostly immunocompetent, whereas infection is fatal in β negative δ negative animals, implying reciprocal compensation by each subset (Sciammas *et al*, 1994). A second example is CMV infection, where recent data highlight a similar scenario, with both the γδ and αβ compartments providing equally potent protection from MCMV infection in the absence of the other (Déchanet-Merville, unpublished observations). Such systems are potentially highly relevant to clinical scenarios involving infection in immunosuppressed individuals, whereby γδ T cells may be able to compensate for a deficiency in the αβ T cell compartment. Two of the most obvious examples include the post-transplant setting, and in neonatal immunity, both scenarios where αβ T cell responses can be highly compromised.

With these scenarios in mind, I will briefly outline the evidence that Vδ2-negative γδ T cells play an important role in immunoprotection from CMV infection in humans.

Strong evidence for involvement of γδ T cells in human CMV comes from the setting of renal transplantation, where immunosuppressed recipients frequently develop acute CMV infection. Déchanet *et al* first reported large expansions of Vδ2-negative γδ T cells in this setting, which correlated with CMV infection, and persisted for years (Déchanet *et al*, 1999a; Déchanet *et al*, 1999b). These involved expansions of $V\delta1^+$ and $V\delta3^+$ cells with evidence for oligoclonal expansions focussed on CDR3 regions of restricted length and bearing specific sequence motifs. In addition, relative to cells in uninfected patients, these populations demonstrated increased expression of activation markers, and expression of other surface markers (CD45RA, downregulation of CD28) consistent with an effector memory phenotype, analogous to $CD8^+$ αβ T cells (Pitard *et al*, 2008). Subsequent studies have strengthened the argument that these cells help control CMV. Expansion of the Vδ2-negative population is associated with resolution of viraemia, consistent with a protective role (Lafarge *et al*, 2001). In addition, Vδ2-negative γδ T cells displayed robust TNFα production and cytotoxicity towards target cells infected with CMV (but not other viruses), and could also limit CMV propagation *in vitro*. Moreover, they frequently expressed homing receptors for the intestine, a key site of CMV infection *in vivo* (Halary *et al*, 2005). Importantly, Vδ2-negative γδ T cells bearing similar phenotypic markers and with equivalent functional abilities exhibited durable expansions outside the transplant setting in healthy CMV^+ individuals, although to more modest levels (Pitard *et al*, 2009). Collectively these data strongly suggest that γδ play an important role in protection from acute CMV infection, particular in the immunosuppressed state. Two independent clinical scenarios confirm this conclusion. Firstly, similar expansions

of CMV-reactive Vδ2-negative γδ T cells have been observed after stem cell transplantation, where patients are also immunosuppressed (Knight *et al*, 2010), and also in patients receiving tyrosine kinase inhibitors, where αβ responses are thought to be impaired (Dr David Lewis, unpublished data). In addition, foetal responses to congenital CMV infection involved major Vδ2-negative γδ T cell expansions with comparable anti-CMV reactivity, and featured a dominant clonotype expressing a public Vγ8Vδ1-TCR, suggestive of antigen-driven expansion (Vermijlen *et al*, 2010). This highlights the probability that Vδ2-negative γδ T cells play a key role in protecting the individual from CMV infection in the immunosuppressed setting, including *in utero*, when conventional adaptive responses are immature. The fact such Vδ2-negative populations exhibit dual reactivity to tumour cells (discussed below) provides important clues as the fundamental role of these cells, and the molecular mechanisms underlying their antigen recognition (Halary *et al*, 2005).

1.7.4 Tumour immunosurveillance and potential application in cancer immunotherapy

Prevalence of γδ T cells at epithelial surfaces, combined with their potent effector responses and the idea that γδ T cells may recognise self encoded stress-induced ligands, suggests the possibility that γδ T cells may play a role in tumour immunosurveillance. Studies in certain mouse models of tumourigenesis support this possibility. Firstly, Girardi *et al* demonstrated that mice lacking γδ T cells were highly susceptible to multiple regimens of cutaneous carcinogenesis (Girardi *et al*, 2001). Interestingly, whereas the presence of αβ T cells either inhibited or promoted tumours depending on the model, γδ T cells universally inhibited tumour development, consistent with a non-redundant contribution of γδ T cells to tumour immunosurveillance. Similarly, a study on colon carcinogenesis in mice detected a higher frequency of carcinoma and dysplastic crypt architecture in TCR δ negative mice, whereas

depletion of the αβ T cell subset had little effect. Again, these results are consistent with a role for γδ T cells in the suppression of epithelial tumourigenesis (Matsuda *et al*, 2001). Finally, CMV-induced expansions of Vδ2-positive cells in renal transplant patients were found to correlate with lower rates of post-transplant skin cancer, consistent with the dual reactivity of such Vδ2-positive γδ T cells to both CMV and tumour cells (discussed below) (Couzi *et al*, 2007).

In keeping with the above findings, there is considerable interest in the use of γδ T cells in immunotherapy approaches targeted at cancer. Several phase 1 or phase 2 clinical trials have been conducted, focusing on exploiting the anti-tumour properties of the Vγ9V2 subset of phosphoantigen-reactive γδ T cells in the context of either haematological malignancies (Wilhelm *et al*, 2003) or solid cancers (Dieli *et al*, 2007; Meraviglia *et al*, 2010). In each case objective clinical responses correlate with the potency of the γδ T cell response. In addition, studies in mice have shown γδ T cells can critically enhance chemotherapy induced anticancer immune responses (Ma *et al*, 2011), suggesting potential for combining chemotherapy and γδ T cell immunotherapy in humans. Finally, the fact that CMV-induced expansions of human Vδ2-negative γδ T cells correlate with lower rates of post-transplant malignancies (Couzi *et al*, 2010) suggest the potential for future vaccination/immunotherapy strategies to enhance γδ T cell-mediated immunoprotection. Together, these studies imply the ability of γδ T cells to respond to tumour cells may be exploited in the clinic in novel cancer immunotherapy approaches.

1.7.5 Immunoregulation and maintenance of tissue integrity

Although mice deficient in γδ T cells do not show obvious signs of autoimmunity or gross immune dysregulation in the steady state, in the context of pathogen infection or challenge with non-microbial stress the same models highlight key roles for γδ T cells in both immunoregulation and in the maintenance of tissue integrity. A good example is the protozoan Eimeria, which infects intestinal epithelial cells. In this system, γδ T cell negative mice exhibit excessive damage to intestinal villi during primary infection, and a heavy neutrophil infiltration, a phenotype reduced by γδ T cell adoptive transfer. Importantly, this excessive inflammation is not observed in TCR β negative δ negative mice, indicating the capacity of γδ T cells to regulate the αβ T cell response that results in immunopathology (Roberts *et al*, 1996). Comparable effects are found in Listeria infection (Thoma-Uszynski *et al*, 1997). Mycobterium tuberculosis infection of mice highlights the potential of γδ T cells to shape the downstream response, as γδ T cell effector responses are required for promoting appropriate cytokine and chemokine signals necessary for protective granuloma formation (D'Souza *et al*, 1997). Beyond infection, an orchitis model demonstrated exaggerated inflammation in the absence of γδ T cells, again suggesting the ability to shape the downstream response (Mukasa *et al*, 1999). In addition, in the MRL lpr mouse model of lupus, γδ T cell deficiency led to increased mortality, correlating with increased autoantibodies, CD4 T cells, and Ig deposition in the kidney (Peng *et al*, 1996). Finally, studies on DETC (Jameson *et al*, 2002) highlighted this population recognizes damaged or stressed keratinocytes, and produces keratinocyte growth factors (KGF) and chemokines. In the absence of DETC, defects in keratinocyte proliferation and tissue epithelialisation are observed, and wound healing in DETC-deficient skin can be restored by adoptive transfer of

DETC or KGF. This suggests a role for DETC in the recognition of injured keratinocytes and subsequent production of factors that directly promote wound repair.

1.8 Molecular basis of γδ T cell regulation

Our molecular understanding of γδ T cell recognition is less advanced than that of αβ T cells. However, I will briefly describe some of the key immune receptors involved in γδ T cell recognition, and summarise our current knowledge of TCR ligands, before focussing on the two populations most relevant to the work outlined subsequently.

1.8.1. Key receptors on the γδ T cell surface

As for αβ T cells, γδ T cell recognition is dependent upon an array of different cell surface receptors, engaging with counter-receptors on the surface of the target cell. Three classes of receptor are worthy of particular attention. Firstly, γδ T cells express a range of NK-like receptors, including both inhibitory (Boullier *et al*, 1998) and activatory receptors (Bauer *et al*, 1999). These include members of the KIR (Lafarge *et al*, 2005), CD94/NKG2, and also activatory NK receptors such as NKp44/46, suggestive of a phenotype intermediate between NK cells and αβ T cells (Pont *et al*, 2012). Most notably, the activatory receptor NKG2D has been highlighted as an important axis influencing γδ T cell activation. NKG2D, expressed on NK cells, CD8 T cells, and γδ T cells, binds to a range of class I MHC-related ligands, in humans comprising MICA, MICB, and ULBP1-6, and in mice RAE1, MULT1, and H60 (O'Callaghan *et al*, 2001; Mistry *et al*, 2007). NKG2D ligands are normally tightly controlled, but can be upregulated by viral infection and in epithelial tumours, and it is notable that both viruses (e.g. CMV) and tumours have evolved ways to evade NKG2D-mediated immune responses (Mistry *et al*, 2007). In the context of tumour immunity, NKG2D negative mice

have defective tumour immunosurveillance in spontaneous models of prostate cancer and lymphoma (Guerra, 2008). In relation to γδ T cells, NKG2D receptor/ligand binding triggers γδ T cell activation and cytotoxicity towards the target cell. Consistent with this, upregulation of Rae1 in the skin initiated altered organisation of DETC population, and resulted in a rapid immunosurveillance response, resulting in infiltration by unconventional αβ T cells (Strid *et al*, 2008). In addition to NK receptor expression, a recent study suggests CMV-reactive γδ T cells commonly express a second class of innate receptor, the Fc receptor CD16 (FcγRIIIA), which is normally associated more closely with NK cells and myeloid cells. CD16 expression enables γδ T cells to produce IFNγ in response to IgG-opsonised virions (Couzi *et al*, 2012). Thus, by expressing NK receptors and Fc receptors, γδ T cells bear some similarities to innate immune subsets.

Despite some evidence for TCR independent effector capacity (Couzi *et al*, 2012; Kapsenberg *et al*, 2009), there is also ample evidence that the TCR is a critical determinant of γδ T cell function. Firstly, transfection of Vγ9Vδ2 TCRs into recipient T cells confers reactivity to relevant phosphoantigens (Bukowski *et al*, 1995). In this subset, NKG2D ligand interactions appear to provide costimulatory signals rather than independently stimulate effect functions (Nedellec *et al*, 2010). In Vδ2-negative T cells reactive to CMV, it appears that NKG2D/ligand interactions do not exert a major role on stimulation, since blocking of either NKG2D or NKG2DL had no effect on target cell recognition. In contrast, anti-CD3 antibodies effectively blocked target cell recognition, and in addition TCR down modulation was observed on antigen recognition (Halary *et al*, 2005). Furthermore, transfection of individual Vδ2-negative TCRs conferred recognition of key target cells (Willcox *et al*, 2012). These data argue strongly that human γδ T cells, comprising both the Vγ9Vδ2 and also Vδ2-negative

subsets, depend heavily on TCR ligation to initiate activation. In mice, recent imaging studies suggest the steady-state contacts between skin-resident DETC γδ T cell population and keratinocytes which take place in intercellular contact zones termed "phosphotyrosine-rich assemblies located on projections" or PALPs, are highly likely to involve TCR/ligand interactions (Chodaczek *et al*, 2012). Moreover, although T10/T22-reactive γδ T cells can survive thymic selection in the absence of the T10/T22 ligands and mature into IL-17-producting effectors, their stimulation once mature is dependent on the TCR (Jensen *et al*, 2008). Therefore, not surprisingly, the γδ TCR, which arose in parallel with the αβ TCR ~450 million years ago, is thought to be a key regulator of γδ T cell function. However, how γδ TCR/ligand interactions take place is largely uncertain.

1.8.2. γδ TCR structures

Given the critical nature of the TCR to γδ T cell function, there are considerable efforts underway to understand the structural basis of γδ TCR recognition. A number of crystallographic studies have shed light on this issue. An initial structure of a human Vγ9Vδ2 TCR outlined some distinctive features of the heterodimer relative to αβ TCR and Fab fragments (such as the interdomain angles and architecture of the C domains), and also highlighted a potential phosphoantigen binding site, comprising an intriguing electropositive patch formed by the CDR loops (Allison *et al*, 2001). Although this has not been verified, this is perhaps not surprising as phosphoantigen recognition is hypothesised to involve an unknown antigen presenting molecule and to involve all CDR loops (Wang *et al*, 2010).

Studies on the mouse non-classical class I MHC molecules T10/T22 system led to a crystal structure of the G8 γδ TCR in complex with T22 (Adams *et al*, 2005). G8 interacted with T22

in an orientation different to that of αβ TCR/MHC interactions, which was intriguing in that involved mainly the CDR3δ loop. A subsequent study extended this mode of recognition to the KN6 clone, and established that grafting of their CDR3δ loops into αβ TCRs was sufficient to confer recognition with wild-type affinity, confirming it was truly CDR3δ-autonomous (Adams *et al*, 2008). While these studies potentially outline an important mode of γδ TCR reactivity, involving only somatically recombined regions (ie CDR3 regions) of the TCR, it remains to be seen how widely applicable this recognition mode is, and clearly involvement of CDR1/2 regions is also possible (Wang *et al*, 2010).

Finally, a recent study details the structure of a human VγVδ1 TCR reported to be specific for the MHC class I related MIC (although these data remain controversial). The structure, which was of the TCR alone rather than in complex with MIC, revealed the CDR loops formed a relatively flat putative binding surface (Xu *et al*, 2011).

1.8.3 γδ TCR ligands and their recognition

The enrichment of γδ T cells at epithelial surfaces suggests the possibility that they act as a first line of defence against epithelial infection, injury or stress. Moreover, based on their restricted V-region usage, one possibility is that γδ T cells primarily engage self-encoded stress ligands, rather than microbial ligands, which are likely to be much more diverse (Janeway *et al*, 1988). However, due to the considerable diversity in their CDR3 regions, the possibility that γδ TCRs recognise foreign antigens to a significant extent cannot be excluded, especially in light of the exclusively CDR3-mediated recognition of T10/T22 by the G8 TCR. Despite ample evidence for TCR dependent recognition of target cells, the ligands bound

directly by γδ TCRs have remained largely unidentified. A summary of putative ligands, in rare cases confirmed by direct binding studies, is given below.

Host-encoded whole protein antigens

Various whole protein antigens have been suggested as direct ligands for the γδ TCR. One of the earliest classes highlighted were heat shock proteins (hsps), which are extremely abundant in the cell, upregulated upon cellular stress, and highly conserved across species. In particular, mycobacterial hsps and GroEL (mycobacterial hsp58/60) were suggested as ligands for both mouse and human γδ TCRs. To date, no direct binding data have been derived that confirm γδ TCR/hsp interaction. In addition, one complicating factor in this area has been the proven ability of hsps to directly bind peptide antigens and deliver it to antigen processing cells (Srivastava, 2002), and the prospect that this kind of mechanism underlies their stimulatory capacity rather than an ability to directly bind the γδ TCR.

MHC and MHC-like proteins

Despite acceptance that γδ T cells do not generally recognise MHC proteins, an alloreactive mouse T cell line, LBK5, has been shown to recognise the class II MHC protein I-E^k (Matis *et al*, 1989; Schild *et al*, 1994; Hampl *et al*, 1999). Mutational studies suggested a binding mode distinct from αβ TCRs that recognise I-E^k (Schild *et al*, 1994; Hampl *et al*, 1999), and hypothesised to be low affinity, although direct binding was not shown. Importantly, the significance of alloreactivity for γδ T cell recognition in general is unclear, but conceivably it could mimic recognition of foreign proteins by γδ T cells (see below). Of relevance to this idea, comparison of γ and δ TCR genes with those of αβ TCRs highlights CDR1 and CDR2

lengths that are longer compared to TCRα and β, and the CDR3s of TCRδ sequences are thought to be more similar to IgH CDR3s (Rock *et al*, 1994). Therefore, instead of the relatively flat surface of the αβ TCR where CDR loop lengths appear to be constrained in order to enable MHC binding, the longer CDRs of γδ TCRs may suggest the ligands recognised are not so constrained.

Despite these arguments, non-classical MHC antigens have long been suspected to be γδ TCR ligands. The limited V-region usage of the γδ T cell population is consistent with a limited number of presenting molecules (Davis and Bjorkman, 1988). Moreover, the ability of γδ T cells to use Vδ regions that can also be used by the αβ repertoire suggests, at least for part of the γδ T cell repertoire, that some MHC-like molecules may be involved – particularly as the Vα domain is thought to play an important role in forming the TCR/MHC interface (Rudolph and Wilson, 2002). The first direct evidence in favour of this hypothesis was the discovery that the mouse non-classical MHC molecules T10 and T22 were ligands were recognised by murine γδ T cell clones (Teitell *et al*, 1994). Although highly related to each other, they are regulated differently – with T22 expressed by most cells, whereas T10 is restricted to splenocytes, thymocytes and macrophages (Ito *et al*, 1990). Notably, neither T10 nor T22 are thought to bind peptide, as the structure of T22 revealed a disrupted binding groove that is too narrow to accommodate peptide antigen (Wingren *et al*, 2000) (see Figures 1.6). Importantly, T22 was the first γδ TCR ligand to be directly confirmed in a biochemical assay (Crowley *et al*, 2000). A tetramer of T22 was found to stain 0.4% of γδ T cells in the blood, and a similar proportion of IEL γδ T cells. BIAcore analysis highlighted an affinity for T22 of 0.1μM, considerably higher than many αβ TCR/MHC interactions. Subsequent structural (Adams *et al*, 2005) and functional (Adams *et al*, 2008) studies have indicated that that the G8/T22

interaction relies entirely on residues in CDR3δ. One of the more important points emerging from these studies is that γδ T cell recognition of T10/T22 is not dependent on antigen presentation, and instead is likely to be affected by increases in T10/T22 expression under certain conditions, such as cellular activation.

Another potentially important group of potential γδ TCR ligands is the CD1 family, of which 5 members exist in humans, CD1a,b, c, d, and e (Brigl, M., and Brenner, M.B. 2004; Willcox *et al*, 2007). CD1 molecules are related to class I MHC proteins in structure, although they possess highly specialized hydrophobic antigen binding grooves, enabling presentation of lipid antigens to T cells, whereby the aliphatic lipid tails are buried inside the CD1 molecules, and the hydrophilic head group is solvent accessible for recognition by the TCR. Whereas CD1d, the only group 2 CD1 molecule, is recognized by Natural Killer T (NKT) cells, the group 1 CD1 molecules comprising CD1a, b and c are known to present a range of pathogen derived (in many cases mycobacterial) lipid antigens (Willcox *et al*, 2007). While the T cells that recognize group 1-CD1-restricted T cells have previously been defined as αβ lineage with diverse TCR V-region usage, there is evidence that γδ T cells may also be involved. Human γδ T cells generated by culturing with CD1-expressing DCs pulsed with mycobacterial extracts were found to recognize CD1c without addition of exogenous antigen (Spada *et al*, 2000). Reactive Vδ1-expressing clones proliferated, secreted cytokines, and killed target CD1c-positive cells. Although responses were blocked by anti-γδ TCR antibody, and CD1c-reactivity could be conferred by transduction of CD1c-specific TCRs, direct recognition of CD1c by the γδ TCR has not been established.

Another set of MHC-like proteins proposed as direct ligands for the γδ TCR are the MHC-class I-related (MIC) genes, comprising MICA and MICB (see Figure 1.6). Studies from the Spies group suggested human intestinal γδ T cells were able to recognise MICA (Groh *et al*, 1998). MICA is attractive as a candidate host-endoded stress antigen, as it is upregulated by heat shock and also during tumourigenesis (Groh *et al*, 1996; Groh *et al*, 1999). In support of these suggestions, $V\delta1^+$ cells from the intestine were found to kill MICA transfectants, and this could be blocked by anti-MIC and anti-TCR antibodies. However, this has remained a highly controversial area, as subsequent work showed MICA/B were ligands for NKG2D, which is expressed by γδ T cells (Bauer *et al*, 1999; Diefenbach *et al*, 2000; Cerwenka *et al*, 2000). Wu *et al*, aimed to resolve this, by transfecting TCR chains from MICA-responsive and non-responsive clones into an NKG2D-negative T cell line (Wu *et al*, 2002). Transfectants that expressed TCRs from MICA-responsive clones were found to bind MICA tetramers effectively, whereas those from unresponsive clones did not. Although detection of direct binding via SPR would provide final confirmation, recent data purporting to highlight this seem unconvincing (Xu *et al*, 2011).

Low molecular weight antigens

The small non-peptidic "phosphoantigens", either host or pathogen-derived, found to activate human peripheral blood subset of γδ T cells are candidate ligands for the Vγ9Vδ2 TCRs expressed by these cells. In support of this, recognition of these compounds can be conferred by TCR transfection, as transfection of a Vγ9Vδ2 TCR into Jurkat T cell line conferred reactivity to isopentenyl pyrophosphate (Morita *et al*, 1995). In addition, reactivity to phosphoantigens has been reported to be dependent on the CDR3 regions of the TCR (Bukowski *et al*, 1998), and was sensitive to small changes in the structure of the antigens

(Morita *et al*, 2001), consistent with TCR-mediated recognition. Moreover, the structure of a Vγ9Vδ2 TCR indicated an electropositive patch, forming a putative phosphoantigen binding site. Despite this, no direct binding data has demonstrated binding of phosphoantigens to the γδ TCR, and it seems likely that an antigen presentation molecule may be involved. In addition to being dependent upon cell-cell contact, presenting cells must be human in origin, consistent with a species-specific antigen presentation molecule (Hewitt *et al*, 2005). Secondly, recent mutagenesis studies highlight all 6 CDR loops are involved in recognition, consistent with a sizeable interface with ligand, and therefore with involvement of an antigen-presentation molecule. Finally, it is difficult to envisage how molecules as small as the phosphoantigens could induce TCR signalling in the absence of a presenting molecule. Clearly further work is needed in this area.

Pathogen-encoded protein antigens

Although γδ T cells use a restricted range of V regions, the high level of diversity in their CDR3 regions suggests that recognition of foreign antigens cannot be discounted. Two lines of evidence suggest this may be the case. Firstly, the TgI4.4 γδ T cell clone, derived from a mouse infected with Herpes Simplex Virus-1, was reported to recognise unprocessed HSV-1 glycoprotein I (gI) (Johnson *et al*, 1992; Sciammas *et al*, 1994). To date, no direct binding data has been obtained to support this. Secondly, recent data from Chien's laboratory has defined both mouse and human cells capable of recognising the model antigen phycoerythrin (Zeng *et al*, 2012). Direct binding studies convincingly showed reactivity was TCR dependent, and comparison of relevant γδ TCRs suggested conserved CDR3 motifs were crucial for recognition, and could mediate recognition in the context of different V regions (Zeng *et al*, 2012). Although this study focuses on phycoerythrin as a model (rather than

physiological) antigen, they demonstrate that γδ TCR recognition of intact pathogen-derived proteins may be a feature of γδ T cell responses.

1.8.4 Molecular basis of ligand recognition by the human Vδ2 negative subset

As noted above, several studies, originally from the laboratory of Dr Julie Déchanet-Merville, have noted substantial and long-lasting expansions of human Vδ2-negative T cells during acute CMV infection, which frequently occurs as a complication of immunosuppression following renal transplantation (Dechanet *et al*, 1999). Importantly, these expansions were in parallel with resolution of CMV viraemia, consistent with a protective antiviral role (Lafarge *et al*, 2001). In addition, they involved multiple V regions and evidence for clonotypic selection, suggestive of antigen-driven expansion *in vivo* (Dechanet *et al*, 1999). In addition, Vδ2-negative T cells induced in CMV infection have an effector memory phenotype (Pitard *et al*, 2008), display strong reactivity to CMV-infected target cells (cytotoxicity, cytokine production) and also were able to limit CMV propagation *in vitro* (Halary *et al*, 2005), supporting the idea they play an important role in anti-CMV responses after renal transplantation.

A landmark publication by Halary *et al* gave important clues to how Vδ2-negative recognition of CMV infection took place. Firstly, use of anti-CD3 blocking antibodies and evidence of TCR downmodulation during recognition strongly supported a key role for the TCR in recognition of CMV-infected target cells. Secondly, CMV-reactive Vδ2-negative T cells were also found to have a dual reactivity, against various epithelial tumour lines. While this was consistent with previous observations that Vδ1-derived cells are able to recognise transformed epithelial cell lines (Groh *et al*, 1998), in this case involvement of NKG2D/ligand interactions

was excluded. These findings, and in particular the dual reactivity towards CMV infection and cellular transformation, suggested the possibility of host-encoded ligands stress ligands that were induced in both infection and transformation. Consistent with this possibility, subsequent work showed that renal transplant patients with CMV-induced expansions has lower rates of post-transplant malignancy, suggestive of an immunosurveillance role for this subset (Couzi *et al*, 2010). Clearly the nature of the targets recognised by the γδ TCR in these responses are of great interest.

1.8.5 Molecular basis of ligand recognition by murine DETC

Although the DETC subset of mouse skin-resident γδ T cells has been discussed in part above, and is also covered more extensively in Chapter 5, included here are a few points regarding how they function at a molecular level. As noted above, studies on DETC have provided strong evidence for roles in immune regulation, maintenance of epithelial integrity, and tumour immunosurveillance. Their epithelial localisation, ability to respond to diverse stress stimuli, and ability to affect downstream responses are characteristic features of γδ T cells in general, and provide strong motivation for understanding these roles at a molecular level. Despite this, and despite the recognised importance of NKG2D/ligand interactions in influencing DETC functions, the exact role of the γδ TCR in these processes, as for Vδ2-negative T cells, is unclear. As outlined more fully in Chapter 5, likely to be of key importance is a molecule termed Skint-1, which is required for thymic development of DETC, and is a candidate ligand expressed constitutively on keratinocytes (Lewis *et al*, 2006) and in the thymus (Boyden *et al*, 2008). Consistent with this possibility, recent imaging data suggest that DETC TCRs may constitutively recognise unknown ligands on keratinocytes even in the absence of stress stimuli (Chodaczek *et al*, 2012). Although these findings highlight Skint-1

as potentially crucial molecule for DETC γδ T cells both in their selection and function in the periphery, our molecular knowledge of Skint-1 and related family members is poor.

1.9 Aims of this thesis

The γδ T cell receptor was identified at approximately the same time as the αβ T cell receptor. Whereas the ligands bound by αβ TCRs are well known, and how they are recognised is well understood, in contrast our knowledge of the ligands recognised by γδ TCRs remains poor. As outlined above, a number of ligands have been proposed, many based on cellular activation assays utilising γδ T cell clones or by use of γδ TCR transfected lines. However in many cases a direct binding interaction between the TCR and ligand has not been convincingly demonstrated.

This thesis aimed to progress our understanding of γδ ligand recognition, and focussed on two systems. Firstly, focussing on Vδ2-negative γδ T cell recognition in the context of CMV infection, I developed a line of research based on identification of Endothelial Protein C Receptor by Dr C. Willcox as a ligand recognised by a Vδ2-negative γδ TCR. This γδ TCR derived from a T cell clone (LES) induced in acute CMV infection that displayed a dual reactivity against CMV-infected target cells and transformed tumour cell lines, properties typical of the Vδ2-negative subset as a whole. Partly to help validate this putative ligand, I first generated soluble γδ TCR for direct binding studies (Chapter 3). To understand the binding mode of the interaction, I carried out an extensive mutagenesis study of the EPCR molecule (Chapter 4). Based on the similarity of EPCR to CD1d, it seemed likely that the LES γδ TCR recognised the α1α2 lipid-binding platform of EPCR, analagous to recognition

of CD1d by the NK TCR. The results had significant implications for understanding how γδ T cell recognition of EPCR might take place in the immune response.

In a second line of investigation, I carried out structural studies on Skint-1, a selecting ligand and putative TCR ligand for the murine DETC subset of γδ cells, as well as a prototypic member of a family of related molecules. In Chapter 5, I describe production of the soluble Skint-1 molecule and both X-ray crystallographic and NMR-based approaches I adopted to define its structure. Following successful determination of a partial structure of the Skint-1 ectodomain, I carried out a comprehensive structural analysis, which is described in Chapter 6.

CHAPTER 2 –

MATERIALS AND METHODS

Chapter 2: Materials and Methods

2.1 Molecular Cloning

2.1.1 Expression Plasmids

pET23a (Novagen) and pGMT7 (Invitrogen) plasmids were used in the prokaryotic expression of Skint-1 and LES γδ TCR. The eukaryotic S2 *Drosophila* plasmid vector, pMT/BiP/V5-His (Invitrogen), was used to express LES γδ TCR, MAU γδ TCR, endothelial protein C receptor (EPCR) and EPCR site directed mutants. The pMT/BiP/V5-His vector contains a metallothionein (MT) promoter for high level, metal-inducible expression of the target gene in S2 cells and a poly-histidine (6xHis) tag for detection and purification of the target protein. The selection plasmid pCoHygro (Invitrogen), consisting of the bacterial hygromycin B phosphotransferase gene under the control of the constitutive *Drosophila* Copia 5′-LTR promoter, was used for stable transformation. The linear maps of plasmids used in this thesis are shown in Figures 2.1 and 2.2.

2.1.2 Primers

Custom primers were designed and then purchased from Invitrogen 50nmol scale. Desalted primers were reconstituted in 10mM Tris pH8/1mM EDTA and stored at -20°C. Primer sequences are listed in Table 2.1. DNA amplification for the gene of interest was performed using proofreading polymerases including Pfu Turbo (Agilent), Phusion (New England Biolabs) and Velocity (Bioline). Polymerase chain reaction (PCR) mixtures typically consisted of 200nM dNTPs (Invitrogen), 20 pmoles of each primer, 1x polymerase buffer, 2.5 units of polymerase enzyme, 200ng of DNA template and made up to 50µl with molecular grade water (Sigma). All thermal cycling experiments were performed using a 2720 thermal

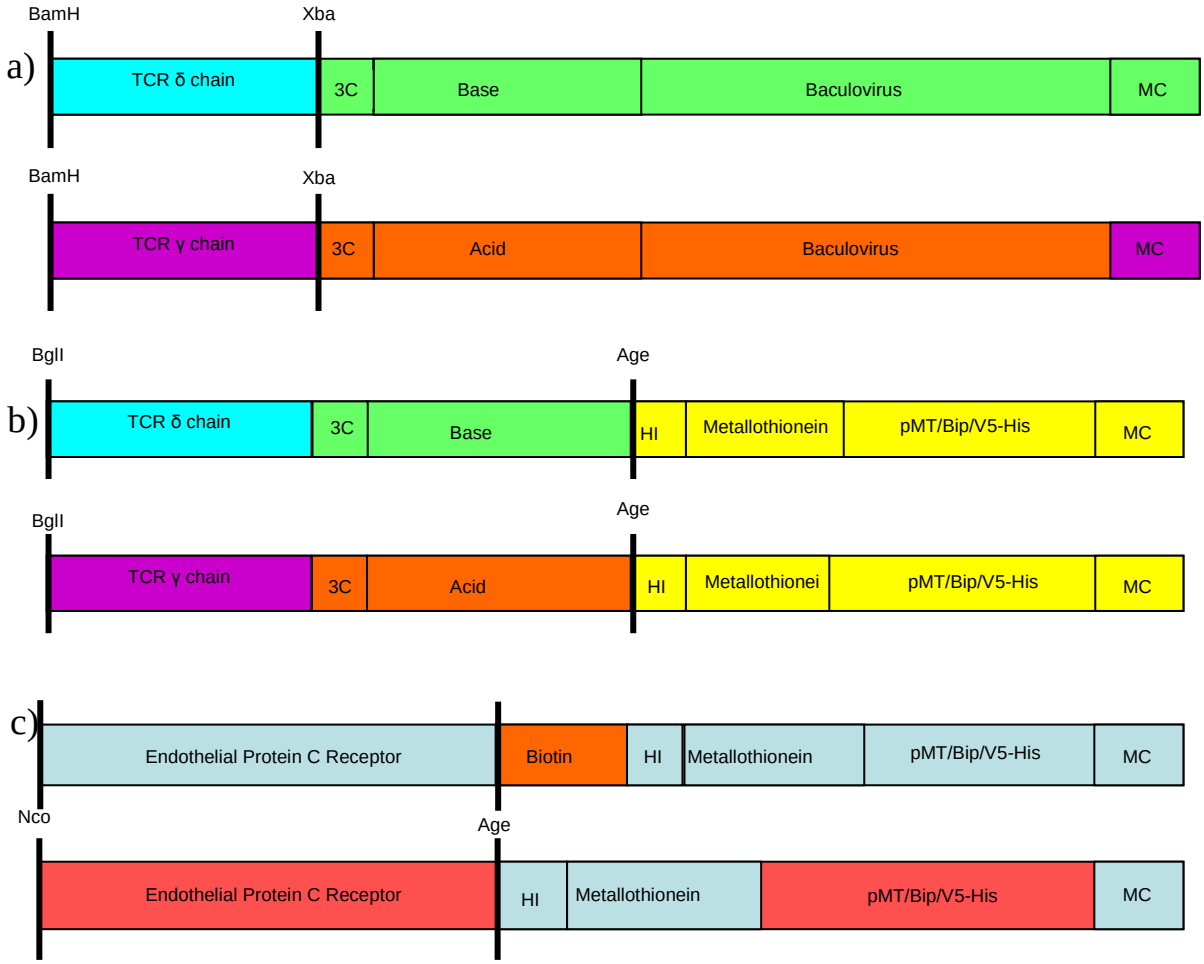


Figure 2.1. Linear plasmid maps highlighting the restriction sites used to clone TCR γ and δ in *Drosophila* (a-b). Linear plasmid map used to express EPCR in *Drosophila* (c).

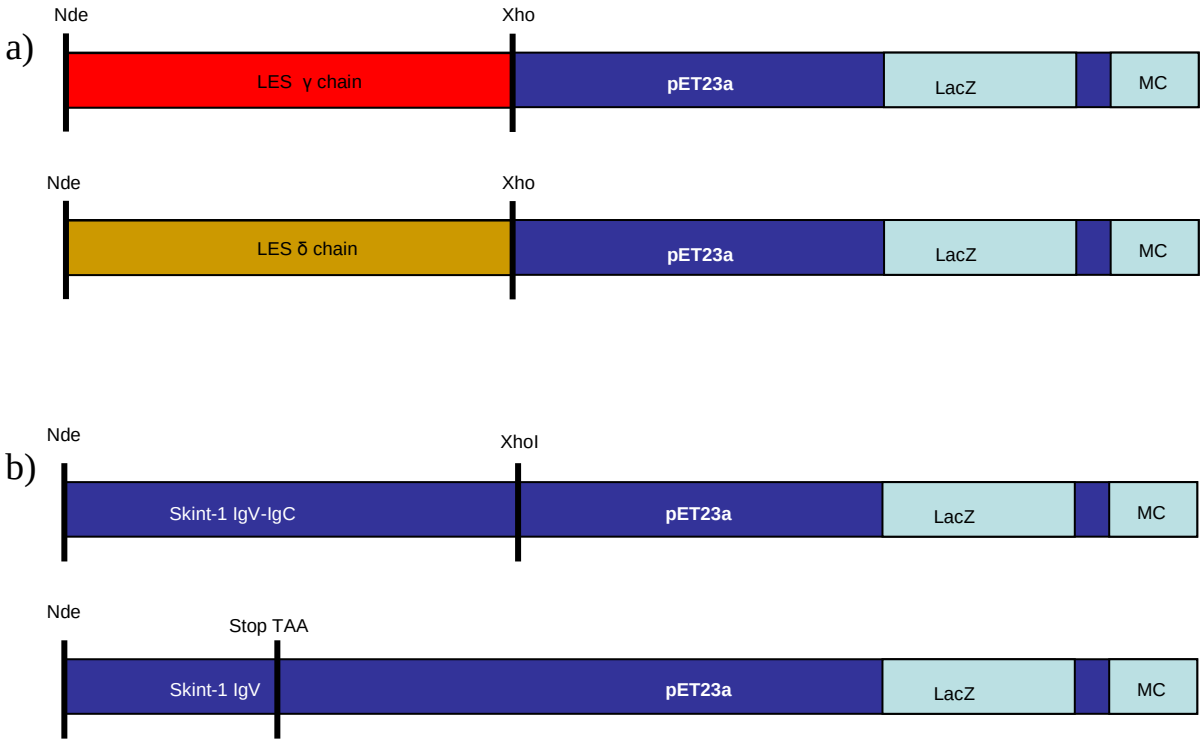


Figure 2.2. Linear pET23a plasmid maps highlighting the restriction sites used to clone LES γ and δ TCR chains (a), Skint-1 IgV-IgC and Skint-1 IgV (b).

cycler (Applied Biosystems). PCR thermal cycling conditions consisted of 95°C for 3 minutes followed by 30 cycles of 95°C for 30 seconds (denaturation), 55°C for 30 seconds (annealing) and 72°C for 1 minute per kilobase (extension), followed by a final extension step of 72°C for 10 minutes.

Primer identification	Sequence
Skint-1 Ectodomain 5’	5’ggg aat tc cat atg agt tca gaa cca ttt ata gtt aat ggt tta gaa ggg ccag tcttggc
Skint-1 Ectodomain 3’	5’ggcctc gag cta ttt cca aat cct att cca tga aaa aaa tg
Skint-1 IgV SDM 5’	5’ cagag gtc aag atc aca tag gcc ata aac tta cagg
Skint-1 IgV SDM 3’	5’ cctgtaagtttatggc cta tgt gat ctt gac ctctg
Skint-1 CCLQ second C SDM to S 5’	5’ cttccagaaagtcatttgttctcttcagaaccctctaac
Skint-1 CCLQ second C SDM to S 3’	5’ gttagagggttctgaagagaacaaatgactttctggaag
LES γ TCR baculo zipper construct 5’	5’ tcccccgggtcttccaacttggaagggagaacgaagtc
LES γ TCR baculo zipper construct 3’	5’ cgtctagagcaactgtctttgggatccactgtggtgacatctg
LES δ TCR baculo zipper construct 5’	5’ cgggatccgaccagcaagttaagcaaaattcacc
LES δ TCR baculo zipper construct 3’	5’ cgtctagagcagctctttgaaggttgctttgtg
LES γ TCR + zipper into *Drosophila* construct 5’	5’gaagatcttcttccaacttggaagggagaacgaagtcagtcatcagg
LES γ TCR + zipper into *Drosophila* construct 3’	5’ ggccaccggtctgagccagttccttttccagtgcttgcaactc
LES δ TCR + zipper into *Drosophila* construct 5’	5’ gaagatctgaccagcaagttaagcaaaattcaccatccctgagc
LES δ TCR + zipper into *Drosophila* construct 3’	5’ggccaccggtctgggcgagtttcttcttgagggcttgaagtttcc
Baculovirus zipper sequencing primer residues E H T S K M	5’ g gaa cac aca agc aag atg
Sequencing forward primer Acid/base zipper residues LLAAAAH	5’ ctt ttg gcg gcg gcg gcg cat tc
Sequencing forward primer Acid/base zipper residues LLAAAAH	5’ ctt ttg gcg gcg gcg gcg cat tc
Sequencing reverse primer Acid/base zipper 3C protease region	5’ cc ctg gaa cag aac ttc cag
Sequencing forward primer Acid/base zipper 3C protease region LEVLFQG	5’ ctg gaa gtt ctg ttc cag gg
Sequencing Basic/Acid Baculo zipper reverse primer QERIRS	5’ c cag gaa agg atc aga tct g
Drosophila primer 5’ N-terminal end	5' caatgtgcatcagttgtg
Drosophila primer 5’ C-terminal end	5' catgtctggatcctcgatc
Reverse TCR δ chain-basic zipper primer to BirA tag	5’ ggcctt ctc gag ctg ggc gag ttt ctt ctt gag g
EPCR 5’ Drosophila primer	5’ ggt ccatgg agccaagacgcctcagatggc
EPCR 3’ Drosophila primer	5’ gg ctcgag cgaagtgtaggagcggcttgtttgg

Table 2.1. Primers used in amplification and sequencing of Skint-1, LES γδ TCR and EPCR.

2.1.3 Restriction enzyme digest

Restriction enzymes (obtained from Roche and New England Biolabs) were used with the appropriate buffer at manufacturer's recommended concentration. Restriction digests were carried at the suggested temperatures for a minimum of 1 hour. To improve downstream experiments many of the digestion experiments were purified using the PCR purification kit (Qiagen).

2.1.4 DNA agarose gel electrophoresis and gel purification

DNA samples prepared in 1x loading dye (40% sucrose, 0.5% bromophenol blue) were analysed by agarose gel electrophoresis at the required percentage (0.6% to 1.2%) run in TBE buffer (89 mM Tris base, 89 mM Boric acid and 2 mM EDTA) at 80V. DNA was visualised by the addition of ethidium bromide. Ethidium bromide intercalates into DNA and fluoresces in ultraviolet light. DNA fragment sizing was carried out by measuring amplified DNA against DNA ladders of the appropriate size (100bp and 1kb ladders; Invitrogen). Agarose gel and solution purification of digested vectors and PCR amplified inserts were carried out with the QIAquick Gel Extraction Kit (Qiagen). Digital images of agarose gels were recorded using the Syngene documentation system and quantified with the Gene Snap software (Syngene).

2.1.5 DNA Ligation

To obtain the most efficient ligation the digested vectors were treated with alkaline phosphatase (Roche) at 37°C for 30 minutes, to prevent self-ligation of the vector and promote the insert DNA ligating into the vector. DNA ligations were carried out at differing

vector to insert ratios overnight using T4 DNA ligase (Roche) as per manufacturer's guidelines.

2.1.6 Lysogeny broth media and pouring LB agar plates

Lysogeny broth (LB) media consisting of 10g bactotryptone, 10g NaCl and 5g of yeast were mixed in a litre of deionised water and autoclaved. For LB agar plates, 15g of agar was added to LB media and autoclaved. The flask was allowed to cool in a water bath (50-55°C) for 30 minutes before adding the antibiotic at the appropriate concentration (e.g. chloroamphenicol 35 μg/ml, kanamycin 50 μg/ml or ampicillin 100 μg/ml). The LB-agar media was poured into plates and left to cool at room temperature for 60 minutes before storing at 4°C.

2.1.7 Transformation of bacteria

TOP10 chemically competent *E. coli* cells (Invitrogen) were thawed on ice. To the thawed *E. coli* cells (50 μl) 10-50ng of plasmid DNA were added and incubated on ice for 30 minutes. After the incubation period cells were heat shocked for 30 seconds at 42°C in a water bath. Immediately following heat shocking, 250 μl of SOC medium was added and then incubated at 37°C for 1 hour at 200 rpm to allow for the resistance gene to be incorporated into the host genome. The final transformation mix was then plated (100 μl) onto a pre-warmed LB ampicillin agar plates (supplemented with appropriate antibiotic) and incubated overnight at 37°C.

2.1.8 DNA miniprep isolation

A single colony from a freshly transformed *E. coli* plate was selected and inoculated into 5ml of LB medium (e.g. 100 μg/ml ampicillin) and grown overnight in an orbital rotating shaker at

37°C/200 rpm. In the following morning, bacterial cell cultures were harvested by centrifugation (13,000 rpm) using a bench top centrifuge (Heraeus Biofuge Fresco 21) to recover the bacterial pellet. Miniprep plasmid isolation was conducted with the QIAprep Spin Miniprep kit (Qiagen) using the manufacturer's guidelines. To obtain higher yields of plasmid DNA, a Maxiprep was undertaken using a Plasmid Maxi kit (Qiagen). The DNA yield and purity were assessed by agarose gel electrophoresis. Successful DNA cloning was confirmed by test digestion of plasmid DNA followed by a final confirmation through DNA sequencing.

2.1.9 Dideoxy terminator sequencing

DNA sequencing was conducted by the Functional Genomics service (University of Birmingham) utilising the Bigdye terminator sequencing kit (Applied Biosystems).

2.2 Prokaryotic protein production

Bacterial protein expression was undertaken using the bacterial transformation protocol as described previously but using a strain of *E. coli* that is modified to express protein (BL21(DE3)pLyS; Novagen), the pLyS is deficient in the ion and ompT proteases which increases the stability of expressed proteins. These cells were transformed with an expression plasmid as described previously and induced by adding 500µM isopropyl β-D-1-thiogalactopyranoside (IPTG) to the culture during the exponential growth phase of the bacteria which then allowed inducible expression of the target protein.

2.2.1 Small-scale test protein expression

Typically, 0.5-1µl of the expression plasmid DNA was transformed overnight into 20 µl of chemically competent BL21(DE3)pLysS *E. coli* cells at 37°C. A single colony was inoculated

into 5 ml of LB culture medium (containing appropriate concentrations of antibiotic). The *E. coli* cells were then cultured by shaking at 37°C until an optical density of 0.4-0.6 OD_{600} was achieved. At this point, a 100 µl aliquot of the pre-induced culture was retained and stored at 4°C. The culture was then induced with IPTG to a final concentration of 0.5mM, after 3-4 hours post-induction, another 100 µl aliquot was retained for comparison with the pre-induced sample using SDS-PAGE. Both samples were subjected to centrifugation for 5 minutes at 13,000 rpm to pellet cells. The pre and post induced pellets were resuspended in 10 µl and 20 µl water, respectively. Reduced loading dye (containing DTT) was added to samples and heated at 95°C for 5 minutes. After heating, samples were vortexed and centrifuged before loading onto the SDS-PAGE gel (8 µl for pre-induced sample and 5 µl for post-induced sample) as described in section 2.3.

2.2.2 Large-scale expression and inclusion body preparation

Large scale expression of protein was conducted in 3-4 litres of LB media cells were grown to O.D. 0.4 to 0.6 and induced with IPTG for 3-4 hours. The bacterial cell pellet was harvested by centrifugation at 5000 rpm for 20 minutes in a Beckman Avanti 25 centrifuge. Harvested cells were resuspended in 50mls of PBS and lysed by sonication with a Misonix 3000 sonicator using 5 x 45s sonication pulses (power level 8) with 15s cooling between pulses. The sonicated preparation was then subjected to centrifugation at 15000rpm for 30 minutes using a Beckman centrifuge (Avanti J-26 XP) to isolate the protein inclusion bodies. The protein inclusion fraction was then subjected to three detergent wash steps to remove contaminants, membrane components and cellular debris. This encompassed homogenisation of the inclusion pellet in 25 ml of Triton containing buffer (500 mM Tris-HCL pH 8, 0.5 % Triton X-100, 200 mM NaCl, 10 mM EDTA, 0.1% (w/v) Na Azide, 2 mM DTT) followed by

centrifugation at 13500 rpm for 15 minutes to recover inclusion body pellet from bacterial lysate. After the three Triton washes, a resuspension wash was conducted (500 mM Tris-HCL pH 8, 200 mM NaCl, 10 mM EDTA, 0.1% (w/v) Na Azide, 2 mM DTT) to remove residual detergent and salt from the protein inclusions. At this point the purified inclusion bodies were solubilised into denaturant buffer (8M Urea, 50 mM MES pH 6.5, 10 mM EDTA, 2 mM DTT) and incubated overnight at 4°C on a 3D gyratory rocking platform shaker (Stuart Scientific). Protein concentration was evaluated using a Bradford dye-binding assay (BioRad). Briefly, 5µl of protein sample (diluted appropriately) was added to 1ml of diluted Bradford reagent (1:5) in a non-UV cuvette and mixed. The spectrophotometer (Eppendorf) was calibrated with a blank and the OD_{595} of the protein sample was measured. The concentration of the test protein sample is determined by comparison against a series of bovine serum albumin (BSA) standards known to reproducibly demonstrate a linear absorbance profile in this assay. The equation below demonstrates how to determine protein concentration.

Protein concentration (µg/ml) = Protein OD reading x dilution factors /0.05 (BSA) standard reading

2.2.3 Oxidative refolding of denatured protein

In vitro refolding of inclusion bodies was performed using the Garboczi oxidative dilution refolding method (Garboczi *et al*, 1992), utilising ~30mg of protein in 500ml refold mixture (100 mM Tris-HCl pH8, 400 mM L-arginine, 2 mM EDTA, 5 mM reduced glutathione, 0.5mM oxidised glutathione and 0.1 mM PMSF) overnight at 4°C. Protein was added in a drop wise manner to reduce protein aggregation. The refolded mixture was then concentrated

to ~10ml using an ultrafiltration unit (Amicon) under nitrogen pressure (50 psi), with an appropriate sized molecular weight cut-off membrane (10 kDa). An alternative oxidative refolding method (Willcox *et al*, 1999) was also attempted to enhance refolding yields. This method used a high concentration of denaturant in the refold mixture (5M Urea, 100mM Tris-HCl pH8, 400mM L-arginine, 2mM EDTA, 5mM reduced glutathione, 0.5mM oxidised glutathione and 0.1mM PMSF) and was left to stir at 4°C overnight. The refolded protein was then dialysed overnight at 4°C to reduce the denaturant levels (1:10 ratio dialysis in 100mM urea). A secondary dialysis step was performed to decrease the concentration of Tris in the refold mixture (1:10 ratio, 100mM Urea, 10mM Tris-HCl pH8 at 4°C). The dialysed protein mixture was then concentrated down and purified by ion exchange (section 2.2.4) and/or size exclusion chromatography (section 2.2.5).

2.2.4 Ion exchange Chromatography

Ion exchange was carried out on a BioCAD®/SPRINT™ Perfusion Chromatography® system (Perseptive Biosystems). The dialysed protein mixture was loaded onto a POROS 10HQ analytical anion exchange columns (Perseptive Biosystems), pre-equilibrated with 0.1M Bis Tris propane pH 8.3 at a flow rate of 5 ml/min. In total, 7 x 200 ml aliquots were loaded, and for each run bound protein was eluted with a linear gradient of NaCl (0-400 mM) over 50 column volumes. Peak eluted fractions were initially analysed by SDS-PAGE, pooled and stored at -20°C.

2.2.5 Size exclusion chromatography

The refold mixture containing renatured protein was filtered using 0.45µm filter disc. The mixture was purified by size exclusion chromatography on a Superdex 200 column pre-

equilibrated with 20mM Tris-HCL pH 8 and 50mM NaCl, using an AKTA FPLC (fast protein liquid chromatography) (GE Healthcare). This purification step was carried out at room temperature at a flow rate of 3 ml/min, and the eluate was collected in 5 ml fractions. The elution profiles were analysed by using the Unicorn (version 4) software (Amersham Biosciences). Fractions corresponding to discrete elution peaks were analysed by SDS-PAGE (section 2.3).

2.2.6 TCA protein precipitation

All experimental procedures were performed on ice unless otherwise stated. To a protein solution, pre-cooled trichloroacetic acid (TCA) was added at a final concentration of 10% (v/v). This solution was mixed and incubated on ice for 5 minutes, after which the sample was centrifuged for 5 minutes at 4°C with a speed of 13000 rpm (Heraus Biofuge). Following centrifugation the supernatant was discarded and the pellet was resuspended with pre-chilled acetone solution. This mixture was centrifuged for 2 minutes at 4°C with a speed of 13000 rpm and the supernatant discarded. The latter step was then repeated and the sample tube was then air dried to remove residual acetone. The pellet was re-suspended with 10-20µls of water and analysed by SDS-PAGE.

2.3 SDS-PAGE

Denaturing protein gels were carried out according to the method described by Schagger and Von Jagow, 1987. Protein purity and quantification was evaluated using the mini protean 3 multi casting chamber system (BioRad). Briefly, SDS-PAGE resolving acrylamide gels of 12-15% (0.765M Tris pH8.8, 0.2% SDS), 5ml 30% acrylamide (Geneflow), 0.1% ammonium persulphate (stock concentration 10%) and 4µl TEMED (Sigma) were mixed and then

pipetted into a 1mm BioRad gel cassette. To level the resolving gel and minimise air bubbles, 1ml of isopropanol was added. Whilst the resolving gel was setting the stacking gel was prepared (150mM Tris pH6.8, 0.12% SDS, 0.1% ammonium persulphate and 4µl TEMED). Once the resolving gel had set the isopropanol was removed and the stacking gel was added on top followed by insertion of a 1.0mm 10-well comb. Protein samples were prepared by mixing ~1-20µl of protein with 2-4µl of 5x loading buffer (containing 125mM Tris pH8, 8M urea, 10% glycerol, 6% SDS, bromophenol blue and +/- 200mM DTT) and made up to a final volume of 10-20µl with deionised water. Samples were vortexed, boiled at 80°C for 5 minutes, centrifuged for 30 seconds before loading onto the SDS-PAGE gel. 10 x SDS-PAGE running buffer (0.1% w/v SDS, 90g Tris, 183g Glycine) was diluted 10-fold and poured into the SDS-PAGE tank. Protein samples were resolved by applying an electric current across the acrylamide gel at constant 100V for electrophoresis through the stacking gel and then subsequently increased to 185V for migrating through the resolving gel. Once the coomassie dye had migrated to the bottom of the gel, the gel was stained in coomassie blue stain (Brilliant blue 0.5g, 250 ml methanol and 50 ml acetic acid) for 30 minutes and de-stained overnight in destain buffer (10% acetic acid, 30% methanol) to visualise protein bands. SDS-PAGE gel images were recorded utilising the Syngene gel documentation system (Syngene).

2.4 Eukaryotic protein production

2.4.1 Expression in S2 *Drosophila* cell line

Protein expression vectors encoding the proteins of interest (pMT/BiP/V5-His) and selection plasmid pCoHygro conferring Hygromycin B resistance were co-transfected into the S2 *Drosophila* cell line. The cells were grown in a selecting medium (Schneider's *Drosophila* medium, modified with L-glutamine, Lonza) consisting of 10% fetal bovine serum

incorporated with 0.3mg/ml Hygromycin B. Every 4-5 days, cells were subcultured into fresh medium containing 5-10% conditioned medium. Small volumes of culture (15-20 ml) were seeded at a density of $1x10^6$cells/ml into T-75 flasks and incubated at 27°C, enabling the cells to grow semi-adherently. For expression, cells were adapted to shaking cultures and were seeded at a density of $1.2x10^6$cells/ml, with 0.05% Pluronic F-68 surfactant (Gibco) to prevent large-scale cell clumping. The flasks were incubated at 27°C at constant shaking rate of 80rpm. Once cells reached log phase protein expression was induced by adding copper sulphate (500 µM) once cells reached a density of $4\text{-}10x10^6$cells/ml. After five days the cell supernatant was harvested by centrifugation (two spins; first at 50 rpm for 10 minutes and second at 2500 rpm for 10 minutes) and following the addition of 0.02 % of Sodium-Azide, stored at 4°C until purification. Expression of *Drosophila* expressed target proteins were analysed by Western Blot.

2.4.2 Western Blot/Dot Blot

Protein samples were first separated by SDS-PAGE as described previously. All separated proteins were transferred (Towbin *et al*, 1979) onto a PVDF membrane (Hybond LFP, GE Healthcare) for one hour (300V/400mA) using a transfer buffer (10% Methanol, 190mM Glycine, 25 mM Tris). The PDVF membrane was incubated for one hour or overnight at 4°C in blocking buffer (5% milk) in order to minimize non-specific antibody binding. The membrane was subsequently washed three times with 20mM Tris pH 7.5, 150mM NaCl (1xTBS) + 0.1 % Tween-20 (Sigma) for five minutes. The membrane was incubated with the primary antibody diluted in 1xTBS + 0.1% Tween + 3% BSA (mouse anti-C terminal HIS (Invitrogen) 1:5000 dilution) for 1 hour at room temperature. The membrane was then washed three times, each for 5 minutes, with 1xTBS + 0.1 % Tween-20. The secondary antibody (HRP-conjugated, anti-mouse, goat mAb 1:20000 dilution, Sigma) was added in 5% milk +

1xTBS + 0.1% Tween and incubated for 1 hour at room temperature. Antibody binding and fluorescence was visualized by using the "EZ-ECL Chemiluminescence Detection Kit for HRP" (Biological Industries). Briefly, this involved mixing 500 µl of solution A with 500 µl solution B and then incubating for five minutes at room-temperature. This mixture was added onto the membrane and incubated for five minutes at room-temperature. The membrane was placed into a cassette and taken to a dark-room, where a light sensitive film (Amersham Hyperfilm^{TM} MP, GE Healthcare) was placed for 10 or 20 seconds on the membrane before the film was developed.

2.4.3 Nickel-NTA Chromatography

S2 supernatant was placed into dialysis tubing (10 kDa cut-off size, Sigma) and dialysed against PBS buffer (1:10 ratio) overnight at 4°C to remove amino acids which compete for binding to nickel beads. 1ml of Nickel-beads (Ni-NTA agarose stored in 20% ethanol, Qiagen, binding capacity of 25mg/ml) was added to an Econo column (BIO-RAD) and equilibrated in 1x PBS + 10 mM Imidazole. The dialysed S2 supernatant was passed over the beads by gravity flow at 4°C overnight at approximately 1ml/min. On the next day, the column was washed with 1x PBS + 10 mM Imidazole to remove the majority of non-specific proteins. The levels of eluted non-specific proteins were evaluated by measuing the optical density of the wash through at OD_{280} at several time points. Once all non-specific proteins had been eluted, the column was filled with 1x PBS + 250 mM Imidazole (~10 ml) and left for 30 minutes. After this period the target His-tagged proteins were collected in 1ml fractions, quantified by measuring optical density at 280 nm and stored at -20°C. Target proteins utilised for binding and structural studies were further purified by size exclusion chromatography.

2.5 Deglycosylation assays

Proteins expressed in *Drosophila* S2 cell lines were utilized in these assays. For analytical samples the deglycosylation experiments with either PNGase F (Roche) or Endonuclease H (New England Biolabs) were performed according to the manufacturer's instructions. Prior to enzymatic treatment, samples were denatured by boiling for 10 minutes in the presence of denaturing buffer (50 mM sodium phosphatase, pH 7.5, 0.02% SDS, 10 mM 2-mercaptoethanol). One unit cleaves 10μgs of protein at 37°C for 1 hour then the cleavage reaction was stopped by heating to 100°C for 5 minutes and deglycoslated protein was evaluated against native protein by SDS-PAGE analysis. For functional experiments, proteins were treated with PNGase F (10ugs protein per unit enzyme) in PBS for one hour at 4°C before surface plasmon resonance analysis.

2.6 3C Cleavage of TCR

Cleavage of purified gamma/delta TCRs, containing a 3C protease cleavage site following the C-terminal cysteine of each chain, was performed using Human RhinoVirus 3C protease enzyme (Novagen) at a concentration of 1U/100μg protein. Cleavage reactions were incubated for 2 hours at 4°C. Removal of excised acid/base zipper and 3C protease enzyme was undertaken by incubating with 200μl of Ni-NTA beads overnight at 4°C with constant mixing. On the following day, confirmation of the cleavage reaction was assessed by SDS-PAGE.

2.7 Surface Plasmon Resonance

2.7.1 Biotinylation of Proteins

Soluble proteins purified by Ni-NTA and/or size exclusion chromatography were buffer exchanged into biotinylation buffer (100mM Tris pH7.5, 20mM NaCl, 5mM $MgCl_2$) using

PD10 desalting columns (GE Healthcare). The PD10 column was initially equilibrated three times with 3.5 ml of biotinylation buffer. Following priming with the biotinylation buffer, soluble protein sample (~3 ml) was loaded onto the PD10 column and then eluted with biotinylation buffer (~3 ml). Biotinylation reactions were carried previously (O'Callaghan *et al*, 1999) using 0.56mM Biotin (Sigma), 1mM ATP (Sigma), 1x protease inhibitors (25x concentrated, Roche) and BirA enzyme (In house product) incubated at 24°C for 1 hour and then at 4°C overnight. The following day excess biotin was removed from the protein sample by size exclusion chromatography as described previously.

2.7.2 Experimental Procedure

Protein-protein binding studies were carried out at 25°C with the University of Birmingham BIAcore 3000 machine (BIAcore, GE Healthcare) using HBS-EP as running buffer (10mM HEPES pH 7.4, 150mM NaCl, 3.4mM EDTA, 0.005% surfactant P20). A new CM5 sensor chip was docked and the system primed with water. Then 500µl of BIAnormalising solution was injected at a flow rate of 10µl/min to normalise the instrument signal. Streptavidin was immobilized onto CM5 binding chips using standard amine coupling chemistry at pH 5.0. Briefly, equal volumes (50µl + 50µl) of 0.2M N-ethyl-N5 (3-diethylaminopropyl)-carbodiimide (EDC) and 0.5M N-hydroxysuccinimide (NHS) were mixed, and immediately injected at 10µl/minute to activate the sensor chip. The next step consisted of injecting 60µl of streptavidin (dissolved in 10mM Sodium Acetate pH 5.5 at final concentration of 0.5 mg/ml) over surface of all four flow cells. This was followed by a third injection with 70µl of 1M ethanolamine pH 8.5 to block the remaining active carboxyl-esters on the surface. Finally 500µl of Glycine-HCl pH2.5 was injected to remove any uncoupled streptavidin. The levels of streptavidin immobilised to each flow cell were calculated in response units, using the

BIAevaluation software. Biotinylated protein samples (in PBS buffer) were diluted in HBS-EP buffer and immobilised to streptavidin at a flow rate of 5µl/min. The remaining flow cells were injected either with control protein, irrelevant protein or left blank. Each flow cell was immobilised with similar levels of protein. Analytes were incubated on ice and centrifuged at 13,000 rpm prior to injections to remove any aggregates formed. Analytes were serially injected over each flow cells at a flow rate of 10µl/minute. For test injections, 30µl was injected. Equilibrium affinity measurements were performed by calculating the analyte/ligand binding response interaction following serial dilutions of analyte (100µM-1.25µM) in HBS-EP at a flow rate of 10µl/minute. The affinity studies were performed in both orientations. Data were collected at 10Hz. Data was analysed with the BIAevaluation software (BIAcore AB) and Origin Pro 8 software (OriginLab Corporation).

2.8 X-ray Crystallography

2.8.1 Crystallisation Trials

Prior to crystallisation trials, protein samples purified by Ni-NTA or size exclusion chromatography were concentrated to 5-10 mg/ml using a microcon centrifugal filter device (Millipore) with a YM-10 cut-off membrane. The purity and homogeneity of the concentrated protein samples were assessed by SDS-PAGE under reducing and non-reducing conditions. Initial crystallisation screening involved using the hanging drop vapour diffusion method with the Mosquito nanolitre crystallisation robot (TTP LabTech). The preliminary crystallisation conditions were identified using commercially available sparse matrix screens from Molecular Dimensions (JCSG$^+$, Pact, MIDAS, ProPlex), Hampton Research (PEG/Ion I and II and Index) and Emerald Biosystems (Wizard I/II and III). The hanging drops were obtained by pipetting 100 nl of protein sample with 100 nl of crystallisation reagent onto a ViewDrop

II 96-well plate seal (TTP LabTech). This was suspended over a 96-well plate (Iwaki) containing 100 µl of reservoir solution, then sealed and allowed to equilibrate by vapour diffusion at 23°C. After 3-4 days each drop was assessed for crystal growth using a light microscope (Leica). The most promising initial crystallisation hits were optimised by varying the pH, incubation temperature, protein concentration, and incorporating additive (Molecular Dimensions) and detergent (Hampton Research) screens. Larger diffraction-grade crystals were grown by manually setting up larger drops using 24-well Linbro plates (Molecular Dimensions). Briefly, equal volumes of protein sample were mixed with crystallisation reagent (usually 1+1µl drops) on a cover slip that was inverted onto a greased well containing 1 ml of the reservoir solution. The drops were then allowed to equilibrate by vapour diffusion for several days before they were viewed for crystal growth as described previously.

2.8.2 X-ray Data Collection and Processing

Prior to data collection, crystals were selected using cryoloops and soaked in reservoir solution supplemented with increasing concentrations of ethylene glycol or glycerol cryoprotectant (5% to 20% in 5 % increments) before being flash-cooled in liquid nitrogen. All crystals were stored in a cryogenic dewar (Molecular Dimensions) prior to X-ray diffraction experiments. All X-ray diffraction studies were carried out using the University of Birmingham Macromolecular X-ray Facility which houses a Micromax 007HF rotating anode X-ray generator (Rigaku) and Saturn charged coupled device. All data collection manipulations were performed using the CrystalClear (Rigaku) software. Intensity data were integrated, scaled and merged using programs of the XDS suite (Kabsch, 2010).

2.8.3 Structure Determination

The determination of initial phase information for Skint-1 was calculated with molecular replacement. Prior to molecular replacement calculations, solvent content calculations were performed with the CCP4 suite to determine the number of molecules per asymmetric unit. Molecular replacement calculations were performed with MOLREP (Winn *et al*, 2011) and Crystallography & NMR System **(**CNS) (Brünger *et al*, 1998). The sequence of the target protein was input into BLAST (basic local assignment search tool) (http://blast.ncbi.nlm.nih.gov) to identify putative search models for molecular replacement.

2.9 Nuclear Magnetic Resonance (NMR)

NMR experiments and analysis were performed in collaboration with Dr Timothy J Knowles using the Henry Wellcome Building for Biomolecular NMR Spectroscopy Facility (University of Birmingham). 1D, 2D and 3D backbone assignment spectra were recorded using Agilent Direct Drive 600, and Inova 800 and 900 MHz spectrometers equipped with HCN 5mm z-PFG cryogenic probes with enhanced 13C, 1H and 15N sensitivity. Machine calibration was performed using the external standard, DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid).

2.9.1 1D Experiments

Unlabelled Skint-1 at a concentration of 0.2mM in 20mM MES, 50mM NaCl pH 6.5 (90%H_2O, 10%D_2O) was analysed in an Agilent direct drive 600MHz spectrometer equipped with HCN 5mm z-PFG cryogenic probes with enhanced ^{13}C, ^{1}H and ^{15}N sensitivity. The standard Biopack (Agilent) 1D water experiment was used, with a pulse width of 8.775 µs,

spectral width of 13008 Hz and 4096 points. The spectrum was Fourier transformed, baseline corrected and analysed using MestraNova (Mestrelab).

2.9.2 Heteronuclear single quantum coherence spectroscopy (HSQC)

^{15}N-labelled Skint-1 at a concentration of 1mM in 20mM MES, 50mM NaCl pH 6.5 (90%H_2O, 10%D_2O) was analysed in an Agilent direct drive 600MHz spectrometer equipped with HCN 5mm z-PFG cryogenic probes with enhanced ^{13}C, ^{1}H and ^{15}N sensitivity. The standard Biopack (Agilent) fast nitrogen HSQC method was used, gNfhsqc, with a pulse width of 9.0 µs, spectral widths of 8397 Hz and 1944 Hz, and 1024 and 128 points in the proton and nitrogen dimensions respectively. The spectrum was apodized, zero filled to 2048 and 256 points, respectively, Fourier transformed, phased, baseline corrected and analysed using NMRPipe software (Delaglio *et al*, 1995).

2.9.3 Hydrogen-deuterium exchange

Hydrogen-deuterium exchange experiments were performed on a 0.2mM ^{15}N-labelled sample of Skint1 in 20mM MES, 50mM NaCl pH 6.5. The sample was divided equally between two standard 5mm NMR tubes (Wilmad), rapidly frozen in liquid nitrogen, then freeze dried using a lyophiliser (Thermo Heto PowerDry LL1500). To one sample H_2O was added to the same volume as the sample prior to freeze drying. This sample was used for HSQC set up, a standard HSQC was taken as reference. Immediately prior to using, D_2O was added to the second sample. Using the calibration performed for the H_2O sample, HSQC experiments were performed within 5 minutes of D_2O addition and then every 10 minutes for 2 hours. Following analysis, those amide resonances remaining after 1 hour were deemed to be involved in hydrogen bonding.

2.9.4 Backbone assignments

Three dimensional backbone experiments were performed on a 1mM sample of $^{13}C^{15}N$-labelled sample of Skint-1 in 20mM MES, 50mM NaCl pH 6.5 in 90% H_2O/10% D_2O. Experiments included standard BEST versions of HNCA, HN(CO)CA, HN(COCA)CB, HNCACB, HNCO, and HN(CA)CO (Biopack, Agilent). All experiments were performed with pulse widths of 9.85 µs, spectral widths in the proton and nitrogen dimensions of 8869 Hz and 2430 Hz respectively at 25^{o}C. HNCA and HN(CO)CA experiments were performed with a spectral width in the carbon dimension of 5174 Hz. HN(COCA)CB and HNCACB, with a spectral width of 13025 Hz, and HNCO and HN(CA)CO experiments with a spectral width of 2454 Hz. Side chain assignments were performed on the same sample using the following experiments: H(C)CH-TOCSY (spectral widths of 6983, 6983 and 5880 Hz respectively in the ^{1}Hx, ^{1}Hy and ^{13}Cz dimensions respectively) , HC(CO)NH (spectral widths of 6757, 8013 and 2000 Hz respectively in the ^{1}Hx, ^{1}Hy and ^{15}Nz dimensions respectively), C(CO)NH (spectral widths of 6757, 7768 and 2000 Hz respectively in the ^{1}Hx, ^{13}Cy and ^{15}Nz dimensions respectively) and HBHA(CO)NH (spectral widths of 6757, 3600 and 2000 Hz respectively in the ^{1}Hx, ^{1}Hy and ^{15}Nz dimensions respectively). NOESY experiments were performed to provide restraints for structural characterisation. The following experiments were performed: ^{15}N-edited NOESY-HSQC (τ_{mix} = 100ms, spectral widths of 8407, 10000 and 2400 Hz respectively in the ^{1}Hx, ^{1}Hy and ^{15}Nz dimensions respectively), ^{13}C-edited NOESY-HSQC (τ_{mix} = 100ms, spectral widths of 9000, 9000 and 7840 Hz respectively in the ^{1}Hx, ^{1}Hy and ^{13}Cz dimensions respectively) and ^{13}C-edited aromatic-NOESY-HSQC (τ_{mix} = 100ms, spectral widths of 7558, 8000 and 8000 Hz respectively in the ^{1}Hx, ^{1}Hy and ^{13}Cz dimensions respectively). All spectra were apodized, zero filled to double the points, phased,

linear predicted, baseline corrected and analysed using NMRPipe (Delaglio *et al*, 1995) and SPARKY 3 (Goddard and Kneller, 2004) software.

2.9.5 Data processing and analysis

Data processing was conducted using the NMRPipe program (Delaglio *et al*, 1995), protein backbone assignments were conducted using the program SPARKY 3 (Goddard and Kneller, 2004). Secondary structure prediction was conducted using backbone assignment data (CA, CB, CO, HA, HN, and N chemical shifts) implemented into the TALOS web server (Cornilescu *et al*, 1999). Proton distance constraints for the structure calculation were obtained through the analysis of three-dimensional ^{15}N-edited and ^{13}C-edited NOESY-HSQC spectra. Cross peaks were manually selected and the peak intensities measured using SPARKY 3 (Goodard and Kneller, 2004).

The 3D structure was determined using the CANDID/CYANA package with the automated NOE cross-peak assignment and structure calculation with torsion-angle dynamics implemented (Güntert, 2004). The CYANA protocol consisted of seven iterative cycles of NOE assignment and structure calculation. The first cycles starts from 200 random conformers with 10,000 torsion-angle dynamic steps performed per conformer in each cycle with the results being used as input in the following cycle. The seven cycles were followed by a final structure calculation where 20 conformers with the lowest CYANA target function were chosen as representative. The final water minimisation was performed with ARIA1.2 (Linge *et al*, 2001). A total of 894 NOES, 38 hydrogen bonds and 182 dihedral-angle restraints were fed in the final calculation. Structures were analysed using PROCHECK-NMR (Laskowski *et al*, 1993). The final structural statistics of the Skint-1 D1 ensemble are

highlighted in Table 4.2 (Chapter 4). The RMSD values from the average structure were calculated using the MOLMOL program (Koradi *et al*, 1996).

Homologous structures were identified using the Dali server (http://ekhidna.biocenter.helsinki.fi/dali). Electrostatic surface maps were calculated with DELPHI (Rocchia *et al*, 2002) and viewed in PYMOL (http://www.pymol.org/). Protein-protein interactions were analysed with programs of the CCP4 suite (http://www.ccp4.ac.uk/). Potential models were generated using one-to-one threading function with PHYRE (www.imperial.ac.uk/PHYRE). Putative protein-ligand docking sites were calculated using the SPPIDER web server (http://sppider.cchmc.org/). All molecular graphic figures were produced with PYMOL.

CHAPTER 3 –

PRODUCTION OF SOLUBLE Vγ2-NEGATIVE γδ T CELL RECEPTOR

Chapter 3: Production of Soluble Vδ2-negative γδ T cell receptor

3.1 Potential utility of soluble receptor for analysing γδ TCR ligand recognition

Studies of αβ T cell recognition have benefited greatly from the availability of soluble T cell receptors (TCRs) and MHC molecules which has enabled analysis of TCR interaction with specific ligand. Early cell-based blocking experiments demonstrated that TCR affinity was in the μM range, several orders of magnitude lower than that of antibodies for their antigens (Matsui *et al*, 1991). Subsequently, production of soluble TCRs has enabled comprehensive biophysical analysis of ligand engagement by the receptor, using various different experimental approaches. Studies carried out using surface plasmon resonance (SPR) to analyse protein-protein binding, as well as structural examination using X-ray crystallography have clarified the molecular features of TCR/peptide-MHC interactions (Garbozci *et al*, 1996; Garcia *et al*, 1996). A wealth of structural studies published over the past 15 years have established a binding mode which places the hyper-variable loops of the CDR3α/β regions of TCR chains directly over the centre of the bound peptide, with CDR1α and CDR1β frequently contributing to peptide recognition by binding over the amino-terminal and carboxyl-terminal peptide regions. The germline encoded CDR1/2 loops of the αβ TCR predominantly contact the alpha helical regions of the MHC molecule (Wucherpfennig, 2010). Direct protein-protein binding assays, chiefly involving SPR; have established the range of TCR/pMHC affinities and indicated TCR binding typically involves slow association/dissociation kinetics and characteristic thermodynamic properties (Garcia *et al*, 1997; Boulter *et al*, 2007). In many cases it is believed these properties reflect conformational flexibility at the TCR binding site, particularly focused on the CDR loops, and are consistent with a model whereby where the unbound state of the receptor is dynamic but stabilised upon binding (Choudhuri K and van

der Merwe, 2007; Willcox *et al*, 1999). Conformational flexibility at the TCR/pMHC interface is likely to increase the potential number of peptide-MHC ligands TCRs bind.

Historically soluble TCR production has been fraught with problems, although several TCR expression systems have emerged. Of these, one method of soluble TCR production involved prokaryotic protein expression and *in vitro* refolding (Figure 3.1) (Garbozci *et al*, 1996). Additional techniques were devised to produce soluble TCR D10 from insect cells (Reinherz *et al*, 1999). The two systems used for insect cell expression are the baculovirus expression method which allows for expression of glycosylated protein in a reasonable quantity. However the gene is not expressed continuously due to the host cell under going lyses and eventual death after infection (Steele *et al*, 2003). The other insect cells expression method incorporates a non lytic expression system where the insect Schneider 2 cells are stably transfected and do not lyse, therefore can undergo several rounds of replication. Although both systems have been used successfully to express TCR proteins, each has advantages and limitations in terms of the quality and quantity of soluble TCR expressed (Garcia *et al*, 1999). Insect cell expression provides folded, glycosylated soluble TCR but in limited yields. Prokaryotic expression can yield the highest levels of protein but the method will require the expressed protein refolded *in vitro* into the native state, a step that is frequently problematic (Hans Peter Sorensen, 2010).

Our understanding of γδ TCR ligand recognition is at an infant stage compared with the field of αβ TCR-peptide MHC recognition. Historically work on αβ TCRs has demonstrated the ligand restriction and the molecular pattern for recognition by αβ TCRs. In contrast, experiments for determining γδ TCRs ligands are still considerably low in numbers.

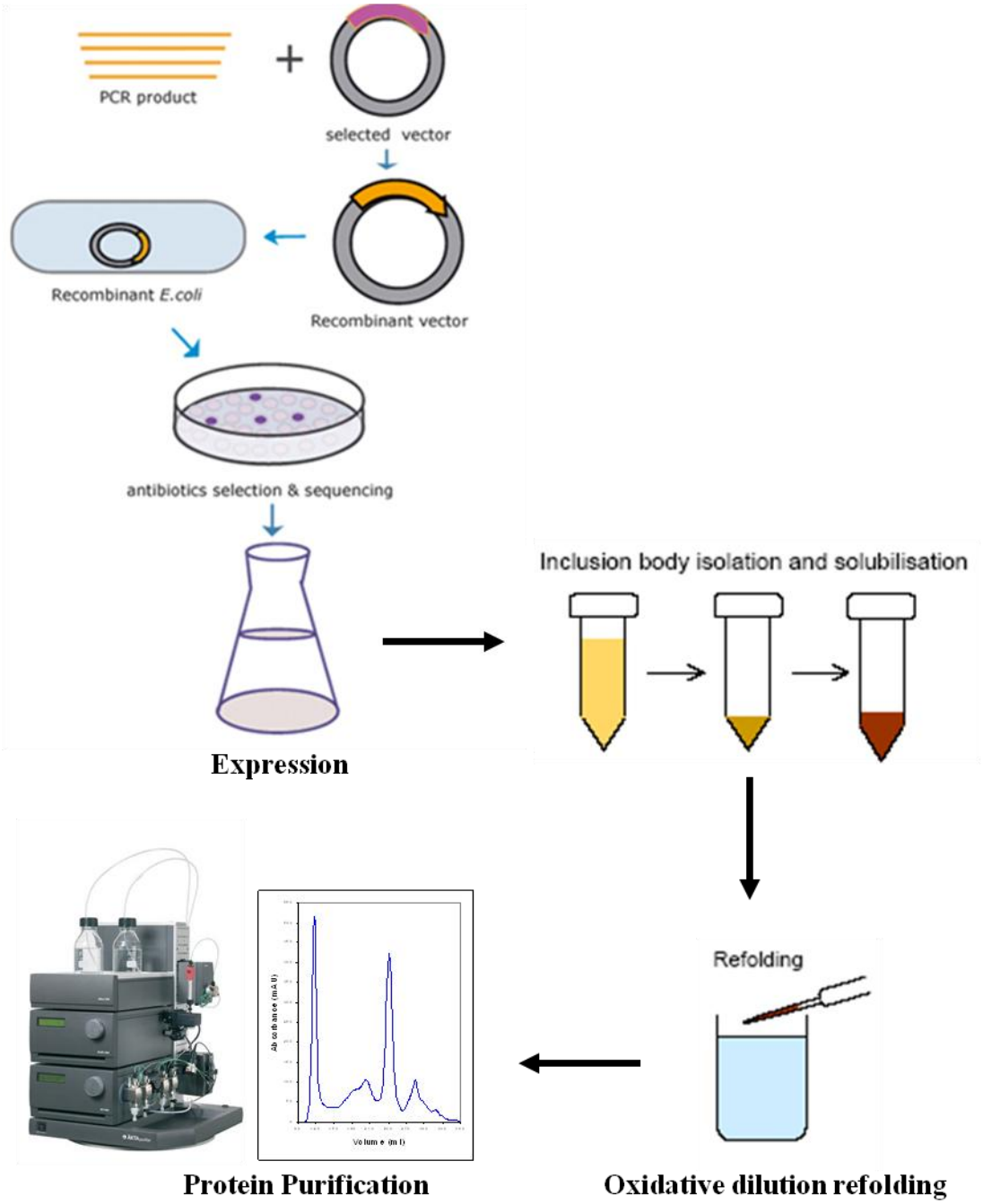


Figure 3.1 Overview of recombinant TCR production in *E. coli*. The γ and δ chains are individually expressed in Rosetta cells as inclusion bodies and renatured by using an oxidative *in vitro* dilution protocol. Following refolding, the mixture is purified by anion exchange and size exclusion chromatography.

Nevertheless the production of soluble γδ TCR will be a valuable tool to help identify key target ligands for relevant γδ TCRs. Two γδ TCRs have been produced and utilised in structural studies. Allison, *et al,* (2001), used a prokaryotic system and then *in vitro* refolding of the TCR. This resulted in the structure of a human γδ TCR. Another technique was published by Crowley *et al,* 2000 using an insect cell expression approach. This technique allowed the characterisation of the interaction for the G8 murine γδ TCR with its non-classical MHC molecule ligands, T10/T22 (Crowley *et al,* 2000) and the solution of the G8 TCR structure in complex with T22 ligand (Adams *et al,* 2005).

3.2 Selection of Vδ2-negative TCRs for soluble production

γδ T cells are known to expand in many different infections and have also been shown to recognise tumour cells lines. Vδ2 negative cells in particular expand in CMV, HIV and other infections (Halary *et al,* 2005; De-Maria *et al,* 1992; Haregewoin *et al,* 1989), and dramatic expansions are often seen in immunosuppressed solid organ transplant patients with acute CMV infection (Déchanet *et al,* 1999). Furthermore individual γδ T cell clones from these patients were shown to have a dual recognition capability by responding to HCMV infected cells and tumour cell lines (Halary *et al,* 2005). The Vδ2-negative γδ T cells responded to HCMV infected cells and were unable to recognise cells that were infected with other viruses. The dual reactivity of these γδ T cells suggested recognition of host ligands induced under conditions of cellular stress, including CMV infection and cell transformation. In particular, two T cell clones were characterised extensively. The LES Vγ4Vδ5 T cell clone was isolated from a patient who received a lung transplant and who developed recurrent acute CMV infections (Lafarge *et al,* 2005), whereas MAU Vγ9Vδ1 T cell was isolated from a renal transplant patient. The TCR genes for these two T cell clones were isolated and transfected

into the TCR deficient jurkat 3 (JRT3) cell line. The JRT3 LES and MAU transfectants were stimulated by the very same cell lines that stimulated the T cell clones (Table 3.1).

Cell Line	LES	MAU
HT29	✓	✓
K562	✓	X
U937	✓	X
HUTU80	X	X
HUH7	X	X
U373	X	✓
293T	X	✓
HeLa EV2	X	X
HeLa PF	X	✓
SKMES	X	✓

Table 3.1. Analysis of different tumour cell lines for γδ TCR recognition. *Cell lines which stimulate the relevant γδ clone or JRT3 transfectant (tick) compared to those that do not stimulate (cross). The LES and MAU γδ clones shows variable suggesting the two clones recognise independent antigens (Halary et al, 2005).*

Crucially, Halary *et al*, (2005) also demonstrated that TCR engagement was an important component of target cell recognition by LES and MAU γδ T cell clones. Use of anti-CD3 and anti-Cδ blocking mAbs to inhibit T cell activation of γδ T cell clones in response to CMV infected foreskin fibroblasts (FSF), suggested a requirement for TCR ligand recognition (Halary *et al*, 2005). Furthermore, specific TCR down regulation was highly suggestive of TCR involvement in T cell activation. In contrast, blocking NKG2D/ligand interactions illustrated that NKG2D was not pivotal in activating the T cell clones. These experiments strongly implied that cell activation was mediated by TCR/cognate antigen recognition (Halary *et al*, 2005). Collectively, these studies suggest that the LES and MAU γδ TCRs

mediate interactions with novel cognate antigens on the surface of tumour/CMV-infected cells, and emphasised these TCRs as interesting candidates for molecular analysis of human γδ TCR. To characterise the molecular interaction binding properties of LES and MAU γδ TCRs, collaboration with the Déchanet-Merville laboratory was undertaken.

3.3 Production of soluble TCR utilising the Prokaryotic expression system

Of the several different soluble TCR expression methods developed, prokaryotic expression systems have been the most widely used. Large numbers of αβ TCRs have been produced from protein purified from *E. coli* cells (Garboczi *et al*, 1996; Simpson *et al*, 2011). Such methods have advantages over insect cell expression in terms of speed, cost-effectiveness and yields (Hans Peter Sorensen, 2010). In addition, *E. coli* expression approaches were successfully used to express, refold and crystallise the G115 Vγ9Vδ2 TCR and determine its structure, providing an important precedent for prokaryotic production of human γδ TCR (Allison *et al*, 2001). Finally, there was considerable experience within the Willcox laboratory with respect to recombinant TCR production using *E. coli* cells and *in vitro* refolding methods (Willcox *et al*, 1999; Nicholls *et al*, 2009; Simpson *et al*, 2011). The strategy devised by Willcox *et al*, for promoting correct refolding of the TCR, consisted of incorporating the heterodimer domain of the c-Jun and v-Fos transcription factors onto the C terminal end of the α and β chains, respectively. Hence during the renaturation process, the Fos and Jun peptides, via their Leucine zipper moieties, facilitated formation of the TCR heterodimer without the requirement of a covalent linkage (Figure 3.2) (Willcox *et al*, 1999).

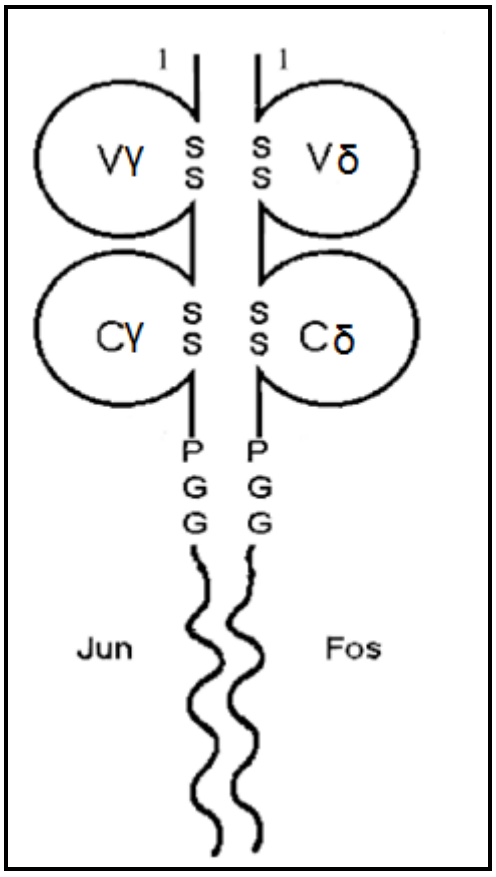


Figure 3.2 Schematic representation of the prokaryotic γδ TCR construct. Leucine zippers derived C-Jun and C-Fos were cloned onto the C terminal regions of the gamma and delta chains, respectively. Leucine moieties were incorporated to facilitate heterodimeric γδ TCR chain association (Willcox *et al*, 1999).

3.3.1 Cloning of LES and MAU TCR Leucine Zipper contructs

In order to employ the Leucine zipper approach to generate LES γδ TCR, the Vγ4 and Vδ5 chains were initially amplified by PCR from plasmid constructs supplied from the Déchanet-Merville laboratory, using 5' and 3' primers enabling cloning into a prokaryotic expression vector, pET23a (see Figure 2.1). Each TCR expression construct contained an optimised 5' region incorporating synonymous nucleotide alterations to improve expression. Pfu amplified PCR products were digested with NdeI and XhoI and purified to enable efficient cloning into pET23a expression vector digested with corresponding restriction enzymes. For the LES Vγ4Vδ5 TCR, the Leucine zipper constructs where cloned using restriction enzymes NdeI and PvuII for the γ chain and NdeI and BglII for the δ chain. A similar cloning strategy was undertaken for the production of the MAU Vγ9Vδ1 TCR zipper chains.

3.3.2 Test Expression and Inclusion body preparation of LES and MAU TCR chains

Having successfully cloned LES TCR zipper chains into pET23a, the Vγ4 and Vδ5 chain constructs were transformed into BL21 (DE3) pLys *E. coli* cells but unfortunately did not over-express protein in sufficient quantities. Hence, Rosetta *E. coli* cells, which are BL21 derivatives and are designed to enhance expression of eukaryotic proteins that possess codons that are rarely used in *E. coli*, were utilised. Over-expression of the LES zipper constructs was verified by SDS-PAGE; test expression experiments indicated a clear induced band for each chain migrating to ~30 kDa, consistent with the predicted sequence derived molecular weight of the γ (~31.6 kDa) and δ (~30.2 kDa) chains, respectively (Figure 3.3a). However, there was a discrepancy in terms of the mobility of the δ chain through the SDS-PAGE gel. Although the predicted sequence derived molecular weight for the δ chain is ~30.2 kDa, an induced band was observed migrating approximately at 34-35 kDa. This deviation from the expected

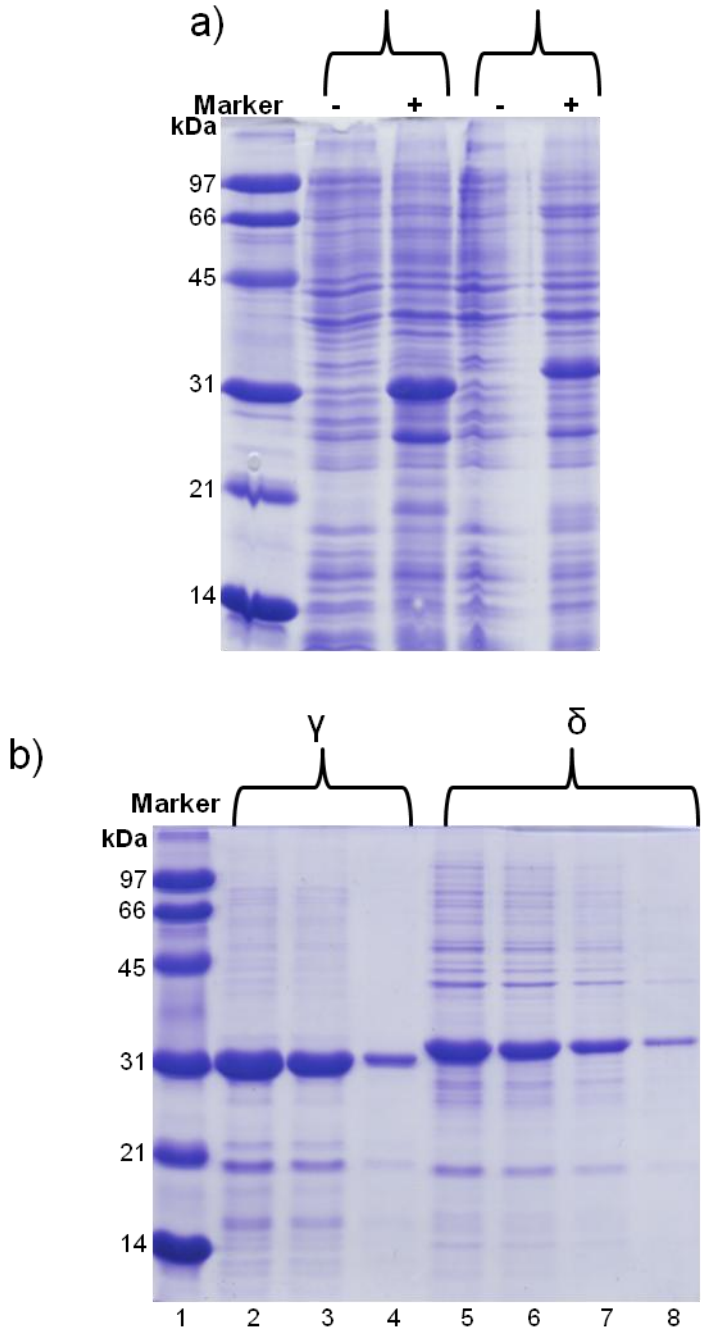


Figure 3.3 a) SDS–PAGE analysis of LES γ and δ TCR chain test expressions in Rosetta *E. coli* cells. Molecular weight standards are shown in Lane 1, with masses indicated in kDa. Lanes 2 and 3 show pre (-) and post (+) induction samples of the γ chain (predicted weight 31.7 kDa), respectively. Lanes 4 and 5 highlight pre (-) and post (+) induction samples of the δ chain (predicted weight 30.2 kDa), respectively. b) SDS-PAGE analysis of LES γ and δ inclusion bodies solubilised in 8M Urea. The δ chain is shown in Lanes 2, 3 and 4 (loaded 13µg, 6.5µg and 0.1µg, respectively), whereas the δ chain is highlighted in Lanes 5, 6, 7 and 8 (loaded 11µg, 5.5µg, 2.75µg and 0.55µg, respectively).

migration on SDS-PAGE could be attributed to the overall fold and charge of the TCR δ chain or due to the number of bound dodecyl sulphate to the protein and finally post translational modifications (Shi *et al*, 2012). Large scale expression and inclusion body preparation of LES γ and δ chains demonstrated that high yields were being produced. For example, typically from a 1L culture, 60-120 mgs of each chain were expressed as inclusion bodies based on the Bradford assay (Figure 3.3b). However, despite such high yields, the purity levels were poor relative to previous *E. coli* derived TCR expressed chains (Nicholls *et al*, 2009; Simpson *et al*, 2011). Indeed, previous studies have highlighted that poor purity of the inclusion body material can adversely affect the refolding process.

MAU γδ TCR chains were also expressed in BL21 cells in this case the MAU γ and δ chains were expressed at high levels. SDS-PAGE analysis of purified inclusion bodies highlighted the MAU γδ TCR chains were considerably more pure (Figure 3.4). Notably, the only difference between the two TCRs is that the Rosetta strain was used for the expression LES γ and δ chains, whereas the BL21 strain were used for the production of the MAU chains.

3.3.3 *In vitro* Refolding and Purification of LES and MAU TCR

Attempts were made to refold equimolar amounts of LES Vγ4 and Vδ5 chains using a previously described method (Allison *et al*, 2001). The refolding buffer consisted of an additive in the form of L-arginine and a strong denaturant 5M urea. In addition, to help form stable disulphide bridges reduced and oxidised glutathione was included into the refolding buffer. After twenty four hours of refolding at 4°C the mixture was dialysed into 100 mM urea for 1 day then a second dialysis into 100 mM urea with 10 mM Tris-HCL pH8 was carried out. The dialysis step aided in generating preferential disulphide bonds that were the

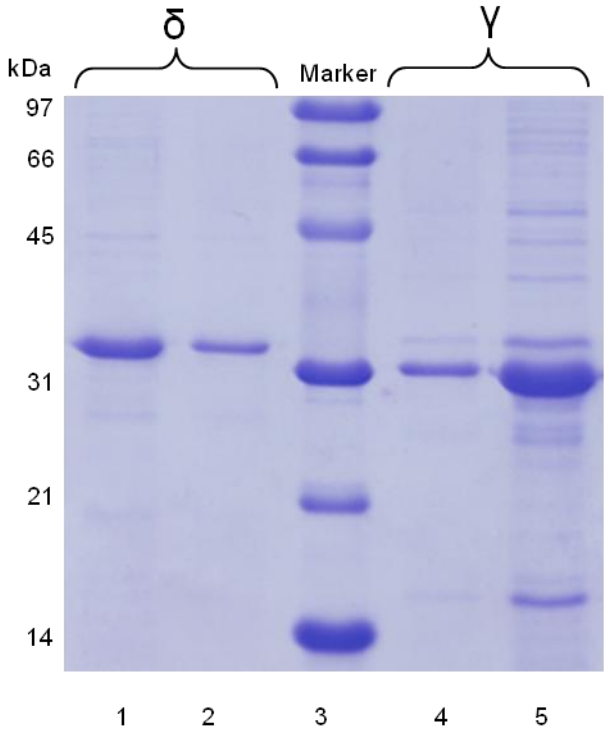


Figure 3.4 SDS–PAGE analysis of MAU γ and δ TCR chains inclusion body expression in BL21(DE3) pLys *E. coli* cells, solubilised in 8M Urea. Molecular weight standards are shown in lane 3, with masses indicated in kDa. lanes 1 and 2 show MAU δ chain (7µg, 1.5µg) (predicted size 32.1 kDa) and lanes 4 and 5 depict MAU γ (3µg, 14µg) (predicted size 33.6 kDa).

most stable. Purification of refold protein was carried out by anion exchange chromatography using a Porous HQ 10 column with a HPLC Perfusion BioCad system. Historically, αβ TCR purified on a Porous HQ 10 column elute as a single peak at approximately 25% of the elution gradient (Nicholls *et al*, 2009; Simpsons *et al*, 2011). However, for the LES γδ TCR the resulting elution profile was vastly different to previous TCR based studies, displaying an extremely broad peak which encompassed the whole elution gradient. Furthermore, the recovery yields for the LES γδ TCR were 90% less than that of αβ TCR using the same refolding strategies (Figure 3.5a). Nevertheless, the eluted fractions were screened by SDS-PAGE to evaluate the quality of refolded protein. From the SDS-PAGE analysis it was clear that both chains were visible under both reducing and non-reducing conditions, with a higher mobility under non-reducing conditions, suggesting formation of internal disulphide bonds (Figure 3.5b). However, the γ and δ chains were non-stoichiometric, as there was generally an excess of δ chain relative to γ chain (Figure 3.5b). Additional purification with size exclusion chromatography provided an elution profile with multiple peaks (Figure 3.6a). SDS-PAGE analysis of fractions corresponding to the elution peak at 190 ml revealed LES γδ TCR, albeit in relatively small yields (Figure 3.6b). The recovery yields of soluble LES γδ TCR were extremely low (less than 2%), and reproducibility of the refolding procedure was poor. Broadly similar results were obtained for the MAU γδ TCR highlighting an extremely low yield (data not shown). Collectively, these refolding experiments indicated the methods developed by Allison *et al* for the production of a Vγ9Vδ2 TCR did not result in robust production of the Vδ2-negative LES and MAU TCR heterodimers.

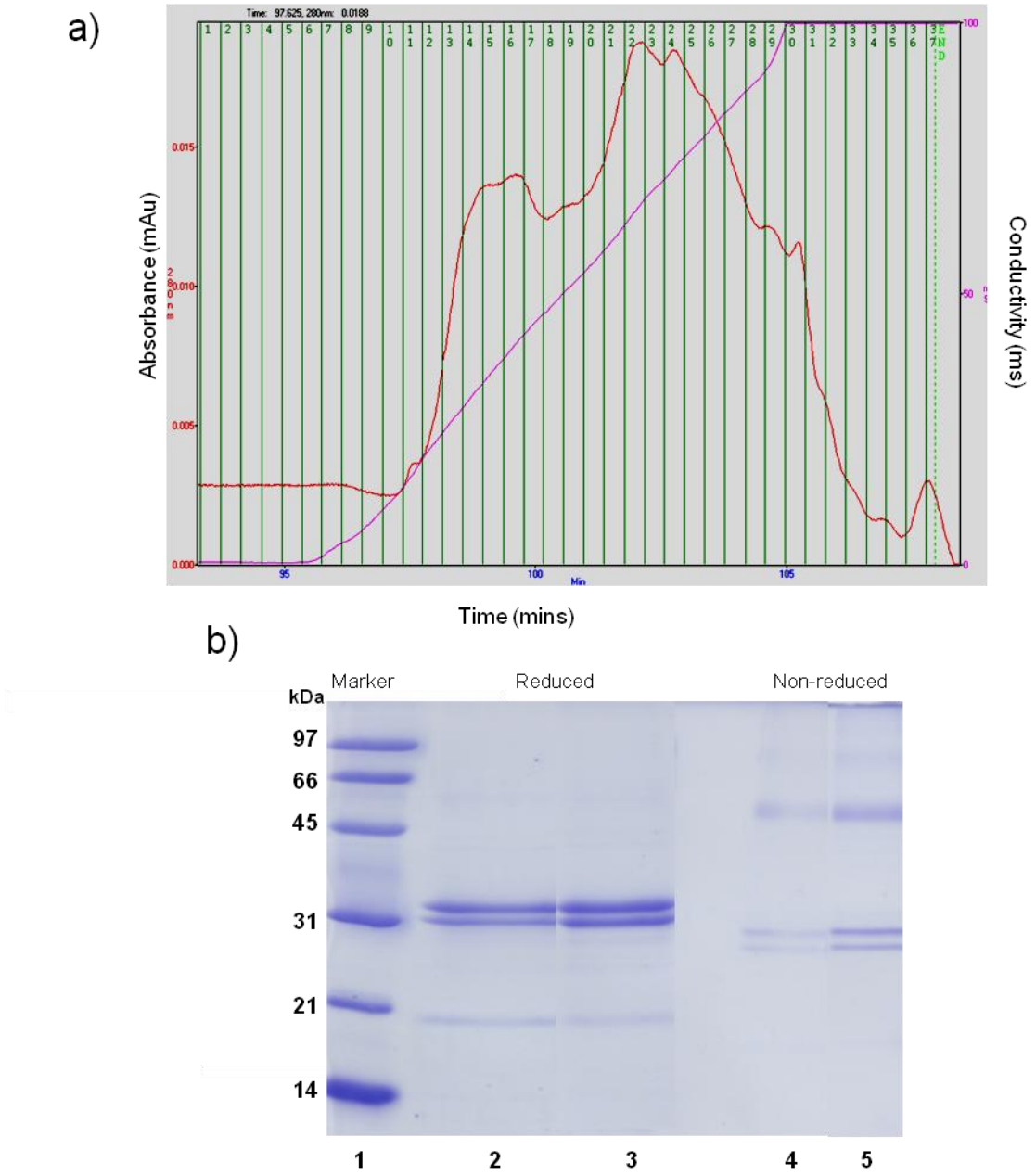


Figure 3.5 a) Anion exchange elution profile for the LES γδ TCR using the 10HQ porous column. Fractions 1-27 (0.5 ml each) are highlighted in green. 200 mls of dialysed LES γδ TCR refold mixture was loaded. b) SDS-PAGE analysis of TCA precipitated fractions 10-23 combined with 25-36 (Lanes 2 and 4) and fraction 24 (Lanes 3 and 5) under reducing and non-reducing conditions. Molecular weight standards are shown in Lane 1, with masses indicated in kDa. The refolded LES γδ TCR is non-stoichiometric.

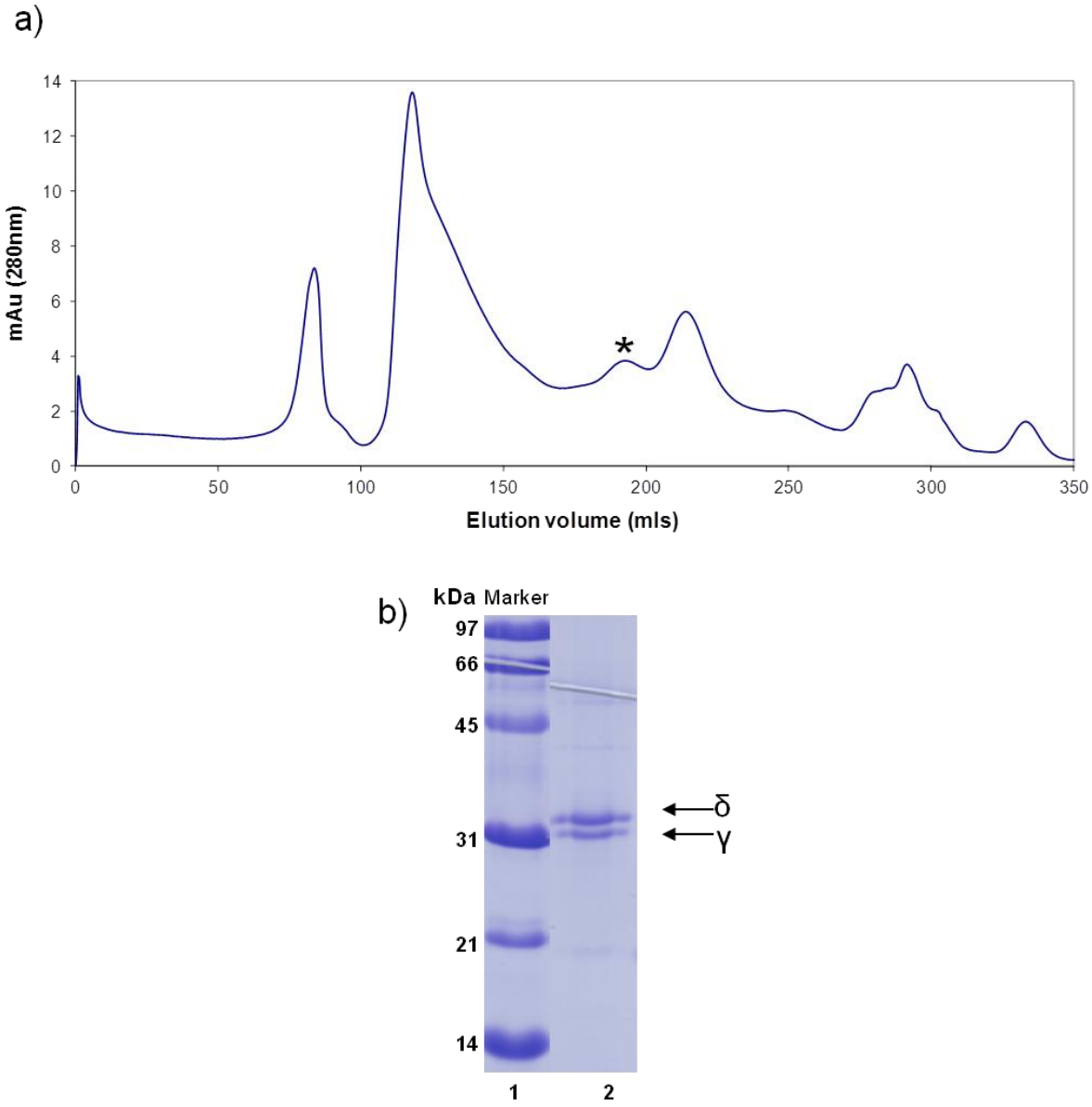


Figure 3.6 a) Size exclusion elution profile of LES γδ TCR following purification using the Superdex 200 column. Fractions eluted from the anion 10HQ anion exchange column were pooled and loaded onto the size exclusion column. The elution peak observed at ~190 ml (~60 kDa) corresponds to LES γδ TCR (shown at *). b) SDS-PAGE analysis of TCA precipitated peak fraction (Lane 2) highlights low level recovery of the LES γδ TCR. Molecular weight standards are shown in Lane 1, with masses indicated in kDa.

3.4 TCR Production using the *Drosophila* expression system

An alternative method to express heterodimeric γδ TCR was required as expression with the prokaryotic system coupled with *in vitro* dilution refolding strategies failed to yield the quantities and purity sufficient for downstream binding and structural studies. This led to attempts at expressing the LES and MAU γδ TCRs using the *Drosophila* cell expression system. The protocol required the transfection of S2 *Drosophila melanogester* cells with both γ and δ TCR chains cloned into separate pMTBiPV5His vectors along with a hygromycin expression vector to select transfectants. The expression system is controlled by the *Drosophila* metallothionein promoter, which allows high level inducible expression of the gene of interest by addition of copper sulphate.

3.4.1 Generating TCR γδ constructs for *Drosophila*

To express protein in the *Drosophila* expression system the cDNA for each TCR chain was cloned separately into the pMTbipV5his vector. Polymerase chain reaction (PCR) of the LES γ chain and δ chain was carried out to amplify the variable and constant region of each chain, enabling cloning into a shuttle vector which contained an acid or base zipper which is approximately 6 kDa in size. The purpose of the zippers was to facilitate correct pairing of the γ and δ chains. Upstream from the acid/base zippers was an addition of a Human Rhinovirus 3C protease cleavage site sequence. This would enable the removal of the zippers after purification. Once correct sequences for the shuttle vectors were confirmed, a secondary PCR was undertaken to amplify the individual chains and corresponding zipper to allow transfer into a S2 *Drosophila* expression vector pMTBiPV5His using restriction sites BglII and AgeI. The S2 expression plasmid vector for each individual chain also contained a V5His tag

sequence to allow purification of the γδ TCR from S2 supernatant by nickel chelate chromatography (For schematic of constructs see Chapter 2 Figure 2.2).

3.4.2 Transfection of LES and MAU γδ TCRs into the *Drosophila* system

In order to express heterodimeric LES γδ protein it was necessary to co-transfect S2 *Drosophila* cells with the expression plasmids containing individual γ and δ chains. Previous evidence has highlighted TCR β chains as having the ability to form β chain homodimers (Garcia *et al,* 1996), therefore it is possible that γ or δ chains could also form homodimers. Hence, single chain transfectants were generated along side γδ co-transfectants. Using the calcium phosphate transfection protocol, each chain was transfected independently or co-transfected with the γ and δ chains as well as the pCoHygro plasmid which contained the hygromycin resistance gene. Transfectants were selected for hygromycin resistance and evaluated for expression of single chains or TCR homodimers following induction by copper sulphate.

Verification of LES γδ TCR expression by the S2 *Drosophila* cells was undertaken. Briefly, the expression levels of LES and MAU γδ TCR heterodimers as well as single chains were examined by western blot analysis of culture supernatant using a mouse anti-His antibody as a probe (Figure 3.7). Western blot analysis demonstrated that LES and MAU γδ TCRs were expressed by S2 cells. Unsurprisingly, both γ and δ chains were required for optimal expression and folding. Dot blot analysis of native culture supernatants using an antibody specific for the acid/base zipper also confirmed TCR chain pairing only in the co-transfected cells (Figure 3.8).

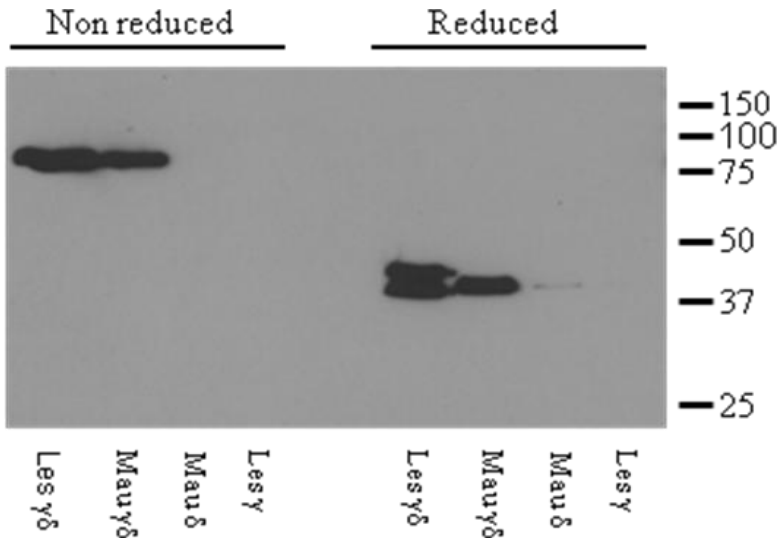


Figure 3.7 Western blot analysis of S2 supernatant following single chain and dual transfections. The blot was probed with an anti-His mAb to validate expression of heterodimeric γδ TCR. LES γ chain displays weak expression relative to the reduced LES γδ transfection which show a double band with increased expression of protein. The non reduced LES γδ band highlights disulphide bond formation between the two TCR chains. MAU indicates similarly high expression levels with dual transfection compared to the single chain transfection.

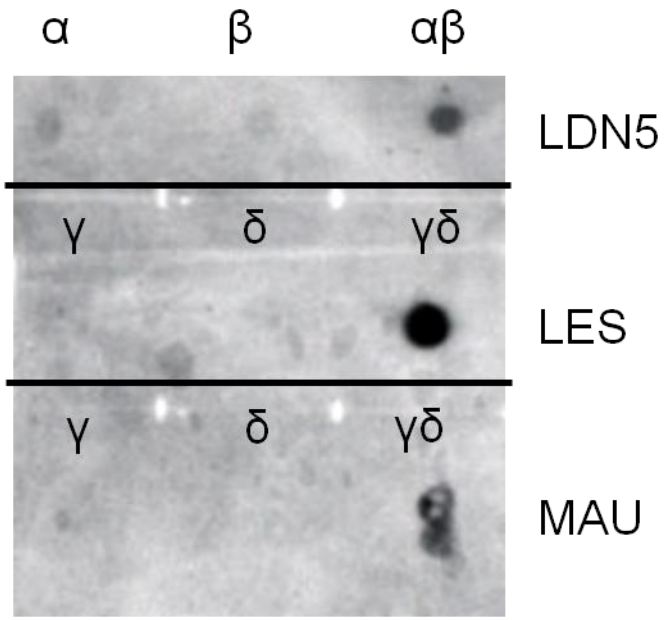


Figure 3.8 Verification of heterodimer TCR expression by dot blot. The dot blots probed with 2H11 antibody demonstrate expression of heterodimeric TCR in cells with dual transfections. S2 cells transfected with single chain highlight no expression.

3.4.3 Expression, optimisation and large scale expression

As previously discussed the expression of LES γδ TCR was verified by western blot analysis using an anti-His antibody. To determine the optimal time point for harvesting TCR following copper sulphate induction, cells were induced and supernatants harvested at the indicated time intervals (Figure 3.9). These time course experiments revealed that the expression levels of protein from S2 *Drosophila* cells remained constant after day three (Figure 3.9). Supernatants were generally harvested on day four after induction.

3.4.4 LES and MAU γδ TCR purification by Ni-NTA chromatography

Recombinant LES and MAU γδ TCRs from S2 *Drosophila* cultures was purified by nickel chelating chromatography. This setup enabled efficient isolation of His-tagged proteins from large volumes of crude S2 culture supernatant. The first step in achieving optimal binding of His-tagged proteins to the Ni-NTA column entailed removing free amino acids such as glycine and glutamine from the supernatant by dialysis against phosphate buffered saline solution stirring overnight at 4°C. The dialysed S2 supernatant was then loaded onto a Ni-NTA column overnight. After washing the column bed with PBS + 10mM imidazole to remove non-specifically bound proteins, the bound TCR was eluted with a higher concentration of imidazole (250 mM). Typically from a 1L S2 culture approximately 3 mgs of LES γδ TCR were generated and 2 mgs MAU γδ TCR. SDS-PAGE analysis of the Ni-NTA eluted LES γδ TCR revealed two prominent bands at approximately 45 kDa, corresponding to the LES γ and δ chains (Figure 3.10). This difference in predicted size and protein mobility visualised on the SDS-PAGE gel could be due to protein glycosylation, as the γ chain has three N-linked glycosylation sites and δ chain has two N-linked glycosylation sites. In addition, there were several higher molecular weight proteins visible suggesting sub-optimal

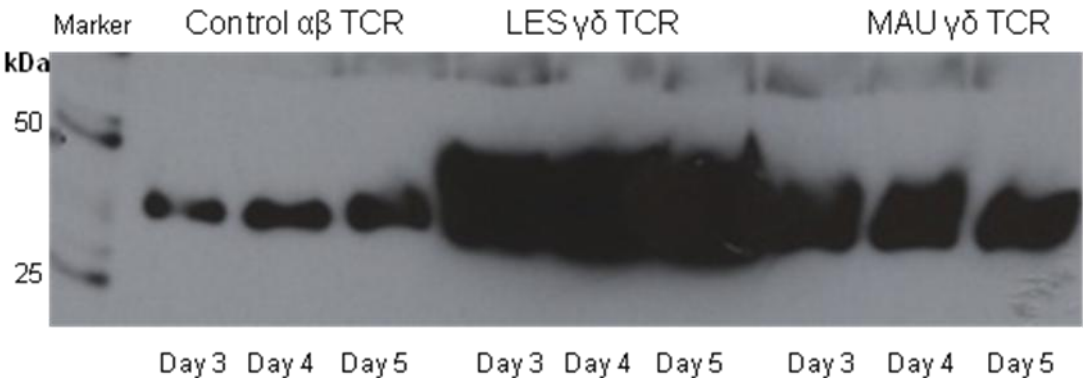


Figure 3.9 Western blot analysis evaluating heterodimeric LES and MAU γδ TCR expression levels at different time points of transfection in S2 *Drosophila* cells. Based on these results day 5 was deemed optimal for harvesting S2 cells. The blot was probed with an anti-His mAb.

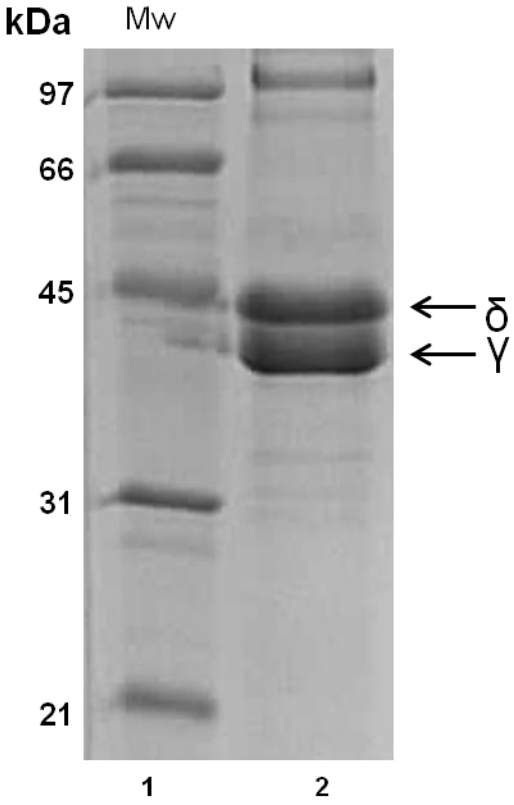


Figure 3.10 SDS-PAGE analysis of purified LES γδ TCR. Molecular weight standards are shown in Lane 1, with masses indicated in kDa. Lane 2 represents LES γδ TCR purified from *Drosophila* supernatant by Ni-NTA chromatography.

purity of the LES γδ TCR following Ni-NTA chromatography (Figure 3.10). MAU γδ TCR demonstrated similar results to LES when analysed by SDS-PAGE (data not shown). Although at this stage the sample was deemed unsuitable for crystallisation trials, the next step involved verifying the correct conformation of the LES and MAU γδ TCR using BIAcore technology.

3.5 Validation of specific heteromeric TCR by BIAcore analysis

SPR was utilised to confirm that LES Vγ4Vδ5 and MAU Vγ9Vδ1 TCRs were conformationaly intact. This was carried out by coupling biotinylated LES and MAU γδ TCRs to a BIAcore CM5 chip coated with streptavidin and injecting a pan-anti-γδ TCR antibody (11F2) over the surface (Figure 3.11). The 11F2 γδ antibody demonstrated specific binding to LES and MAU γδ TCRs relative to the control αβ TCR that was also expressed in the *Drosophila* expression system. In addition, injection of anti-Vδ1 monoclonal antibody over the surface of a flow cell immobilised with LES and MAU TCR demonstrated specific binding to MAU Vγ9Vδ1 TCR relative to LES Vγ4Vδ5 TCR (Figure 3.12a). Also, injection of anti-Vγ9 monoclonal antibody over the surface of a flow cell immobilised with LES and MAU TCR highlighted specific binding to MAU γδ TCR in comparison to LES γδ TCR (Figure 3.12b). At this time point in the experiments there was no specific antibodies against Vγ4 or Vδ5 to further validate the confirmation of the LES γδ TCR. However, utilising the LES γδ TCR the collaborators in the Déchanet-Merville laboratory were able to derive a LES specific antibody. This was later used in SPR experiments to validate the LES γδ TCR by Dr Carrie Willcox (data not shown). Collectively, these observations confirmed the correct association of LES and MAU γδ TCR chains as well as validating the functionality of recombinant TCRs generated by the *Drosophila* expression system *in vitro*. Indeed, the MAU

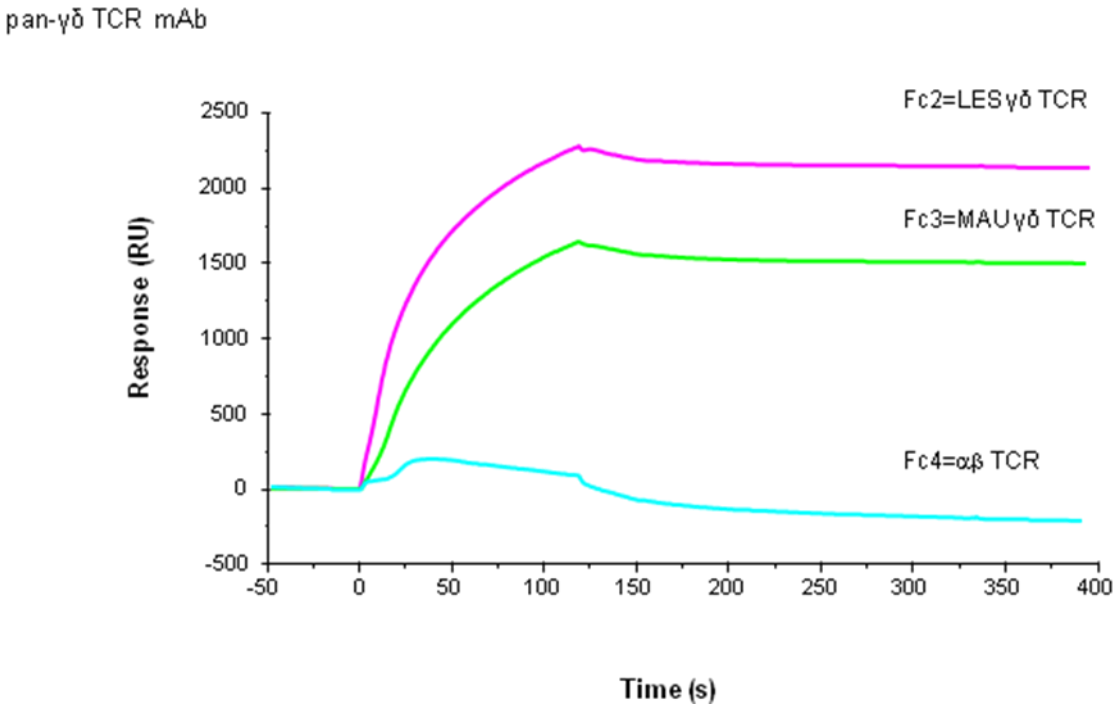


Figure 3.11 Verification of the correct conformation of LES and MAU γδ TCR using surface plasmon resonance. The pan-TCR γδ antibody, 11F2, was injected over the surface of flow cells immobilised with LES γδ TCR, MAU γδ TCR and control αβ TCR. The pan-TCR γδ antibody, 11F2, demonstrates specific binding to LES and MAU γδ TCRs relative to the control αβ TCR.

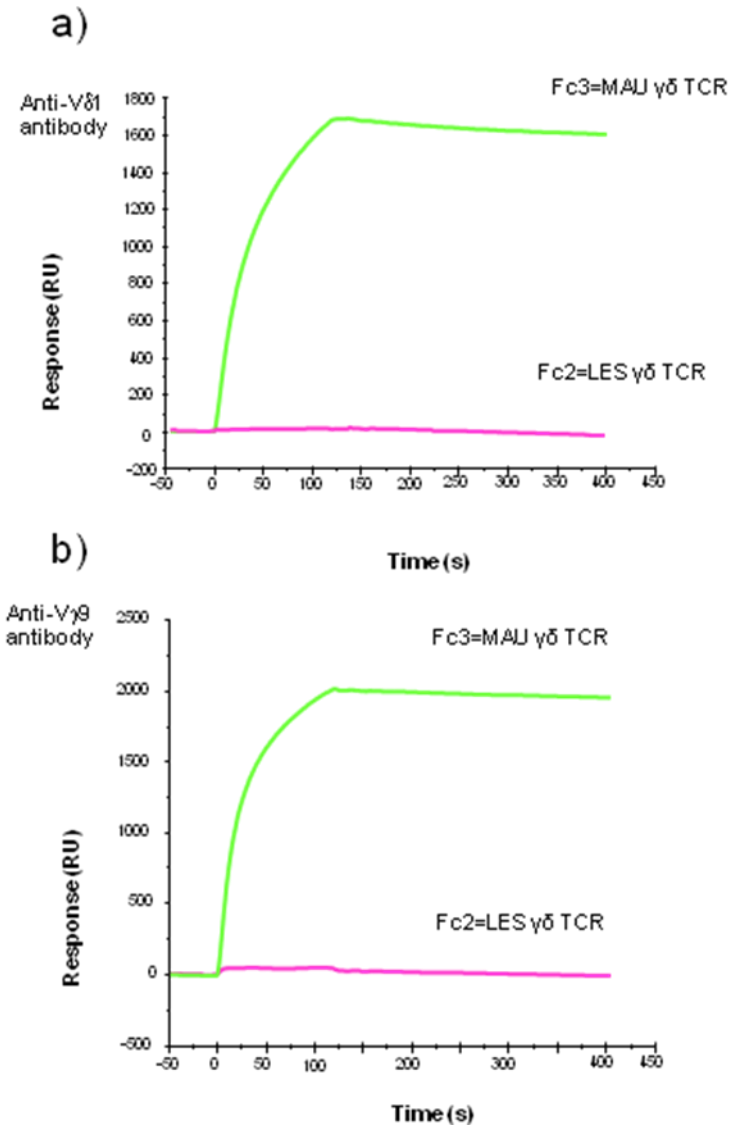


Figure 3.12 Verification of the correct MAU TCR chain conformation using surface plasmon resonance. a) Injection of anti-Vδ1 monoclonal antibody over the surface of a flow cell immobilised with LES and MAU TCR. The anti-Vδ1 antibody demonstrates specific binding to MAU TCR (Vγ9/Vδ1) relative to LES TCR (Vγ4/Vδ5). b) Injection of anti-Vγ9 monoclonal antibody over the surface of a flow cell immobilised with LES and MAU TCR. The anti-Vγ9 antibody demonstrates specific binding to MAU TCR (Vγ9/Vδ1) relative to LES TCR (Vγ4/Vδ5).

γδ TCR was used for downstream ligand identification experiments together with characterising/validating candidate MAU TCR/ligand interactions using SPR and alpha screen technology. Moreover, the LES γδ TCR was utilised for functional (see Chapter 4) and structural studies (see sections 3.6 and 3.7).

3.6 Structural studies of LES γδ TCR with X-ray Crystallography

3.6.1 Overview of X-ray crystallography

X-ray crystallography is a powerful technique used to determine the three dimensional structures of proteins. Structural information can highlight how protein molecules bind to and recognise their ligands as well as to ascertain their function. This technique, however, is heavily dependent on the protein of interest crystallising and diffracting X-rays to a sufficient resolution such that significant structural information can be gleaned (Figure 3.13). In addition, the quality of structural data that can be extracted from a protein structure is extremely reliant on the quality of the crystal.

3.6.1.1 Protein Expression and Purification

For maximal chances of crystallisation, usually high concentrations of relatively pure protein are required. Hence the protein of interest is introduced into an optimal expression system (Prokaryotic or Eukaryotic) that allows high quantities of protein to be expressed. Following expression, conventional protein purification approaches are undertaken, which involves separating proteins based on their differences in charge, size or whether they possess a specific tag (e.g. poly-His tag, GST). The sample is then subjected to a series of preliminary assays to evaluate protein purity and homogeneity using SDS-PAGE and dynamic light

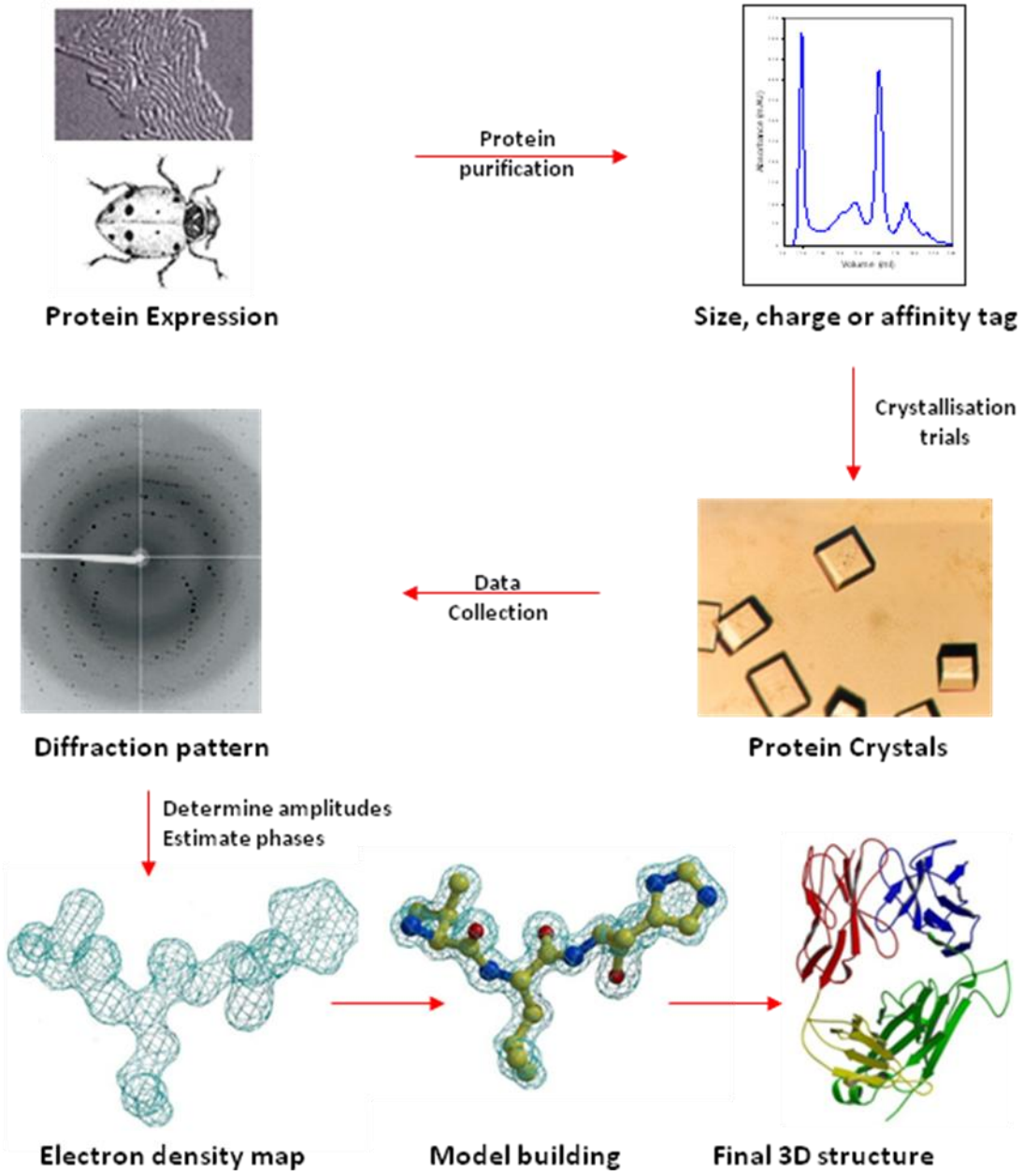


Figure 3.13 Overview of X-ray crystallography. The target protein is over-expressed in a suitable expression system and purified using conventional methods to allow the initial crystallisation trials. Once crystals of a suitable size are grown they are exposed to an X-ray beam and the resulting diffracted X-rays are recorded on a detector in the form of spots or reflections. From the diffraction pattern the data is processed to provide information about the unit cell, space group. In addition, the intensity of each reflection is also measured and converted to amplitude. The phase for each reflection is estimated and when combined with amplitudes an electron density map is calculated. The electron density map is a map showing the distribution of electrons and can be interpreted to enable the molecular structure to be modelled.

scattering. It is important that the protein is devoid of contaminants and exists as a homogeneic species otherwise it can severely hinder the subsequent crystallisation process.

3.6.1.2 Protein Crystallisation theory and practice

Once the protein has been fully evaluated and fulfils the recommended criteria for successful crystallisation, the protein is concentrated to high levels and subjected to crystallisation trials. Crystals are grown by bringing the protein solution to a supersaturated state. This process occurs in three key stages. Firstly, is the nucleation phase in which well ordered aggregates called nuclei (micro crystals) are formed. Secondly, with the growth phase, protein molecules are recruited to the growing faces of the crystal. Finally, in termination phase, cessation of crystal growth occurs due to growth defects, presence of contamination or protein depletion. This supersaturated state forces protein molecules in solution to interact with each other in a specific manner which initiates nucleation and crystal formation, achieved through reducing the solubility of the protein (Li and Ismagilov, 2010).

There are several factors that influence the solubility of a protein. One of the most important parameters is pH, since the solubility of a protein is at the lowest at the isoelectric point (i.e. the pH at which the protein has no net overall charge). Another key factor is ionic strength as introduction of salts (e.g. sodium chloride) have a dehydrating effect on the protein as it competes with protein molecules for water. Moreover, polymers such as poly ethylene glycol reduce the dielectric constant of the solution as well restructuring the drop by volume exclusion which in turn forces protein molecules to interact with each other. Collectively, the overall crystallisation strategy with the use of buffers, salts and precipitants is to reduce protein-solvent interactions whilst enhancing protein-protein interactions (Bergfors T, 2007).

These are commercially available as screens. The most common method used to crystallise proteins is the hanging drop vapour diffusion method. This involves mixing equal volumes of protein and crystallisation reagent onto a cover slip, which is inverted on a greased well containing the crystallisation reagent. The hanging drop is sealed to form an enclosed system. By creating differences in crystallisation reagent concentration between the drop and the reservoir solution, water molecules diffuse out of the hanging drop by a process called vapour diffusion until equilibrium is achieved. During this process, the protein reaches a supersaturated state and protein molecules are forced to interact with one another and if this interaction is specific, nucleation begins and then leads to crystal growth. The identification of crystallisation conditions optimal for crystal growth is largely a multi-dimensional search by trial and error and often represents one of the biggest obstacles in X-ray crystallography.

Once a preliminary crystallisation hit has been identified, it is optimised to grow larger crystals for subsequent X-ray diffraction experiments. Prior to this the crystal is selected and soaked into the reservoir solution supplemented with cryoprotectant for two reasons. Firstly, it serves as anti-freeze and prevents the solvent surrounding the protein from forming ice crystals and secondly protects the crystal from X-ray radiation damage.

3.6.1.3 X-ray Diffraction Theory and Data Collection

Since individual atoms in protein molecules typically span a range from 0.01-10 nm, they can only be visualised with X-rays as they fall within a similar wavelength (0.01-10 nm range). The crystal is exposed to an X-ray beam, and the electrons surrounding the protein atoms within the crystal diffract X-rays that are recorded by an X-ray detector in the form of a

distinct array of black spots called a diffraction pattern. From the diffraction pattern it is possible to extract information that is necessary for determining the protein structure.

Each diffracted X-ray represents a simple wave that consists of three parameters (frequency, amplitude and phase), and when defined, it can lead to determining the three dimensional structure of the protein. The frequency of the wave can be calculated from the wavelength of the X-ray source used to conduct the X-ray diffraction experiment. The amplitude of the wave specifies the height of the crests and troughs and can be determined by measuring the intensity of each spot and converting to amplitude. The phase of each wave specifies the position of the wave with respect the origin of the co-ordinate system on which the wave is plotted. During X-ray data processing it is possible to measure the amplitude of each reflection but not the phase information and this is in crystallographic terms is referred to as the ‘phase problem’.

3.6.1.4 Techniques for Calculating Phase Information

There are several methods available for calculating initial phase information, but the most common method used is molecular replacement (Evans and McCoy, 2008). With molecular replacement phases from a known structural homologue (e.g. phasing model) are used as initial estimates for the target structure. Importantly, since phases are sensitive to position of the atoms, before meaningful phase information can be extracted, the phasing model is initially superimposed onto the target structural data. This superimposition is heavily dependent on the sequence/structural homology between the phasing model and the target structure. The more homologous the phasing model, the easier it is to superimpose the two, enabling the extraction of phase information and solving the target structure.

However, if there is no structural homologue available, as is the case for most novel structures, phases must be determined experimentally. One method used for this purpose is multiple isomorphous replacement (MIR) and this relies on finding the positions of a few heavy atoms (atoms that are electron dense). Compounds that contain heavy atoms have a significantly higher atomic scattering factor because they possess more electrons in their outer shells which can interfere with the X-rays and this in turn affects the intensity of the reflections. The scattering contribution of these heavy atoms in combination with the knowledge of their positions can be used to calculate phases. The first step involves crystallising the target protein (native) and collecting a complete X-ray data. The native crystals are then soaked in solutions containing heavy atom compounds to generate derivatised crystals. The idea is that the heavy atoms will soak into crystal via the solvent channels and bind to specific amino acids of the target protein crystal. Importantly, the type of heavy atoms used is dependent on the amino acid sequence of the target protein. For example, mercury (Hg), gold (Au) and platinum (Pt) are highly reactive towards Cysteine, Histidine and Methionine residues. In contrast, heavy atoms such as the lanthanides and actinides bind to negatively charged residues such as glutamic acid and aspartic acid (Blundell and Johnson, 1976).

The native crystal is soaked in solution containing the heavy atom compound to generate an isomorphous derivative protein crystal (e.g. the unit cell parameters of the derivatised crystal must remain similar to the native crystal). This can be very difficult to achieve as often crystals soaked in heavy atoms can lead to crystal cracking or dissolving. If the crystal survives the heavy atom soaking process, the derivatised crystal is exposed to X-rays and an X-ray data set is collected. By subtracting the derivatised data from the native data, it is

possible to find the positions of the heavy atoms and this in turn can lead to determining initial experimental phase information. It is important to highlight that by generating a second isomorphous derivative crystal using a different heavy atom, it is possible to calculate more accurate experimental phase information. The advantage of using MIR is that it is a powerful tool for generating phases of novel structures. The disadvantages include that it is extremely time consuming, the user is exposed to toxic compounds and finally there is no guarantee that heavy atoms will bind the crystal after soaking.

The final technique used to calculate experimental phases is multi-wavelength anomalous dispersion (MAD). At the absorption edge of an element X-rays are absorbed and re-emitted with altered phase (anomalous dispersion). These altered phases can provide useful phase information that can be used to solve structures (Hendrickson WA, 1991). The most common heavy atom used for MAD is selenium which is incorporated into the target protein prior to crystallisation. This can be achieved by transforming target plasmid cDNA into methionine auxotrophic strain and then growing the cells in presence of selenium (selenium substitutes for methionine). Notably, with this approach the heavy atom is already incorporated into the target protein thus relieving the need for preparing multiple derivatised crystals as in MIR. The seleno-Met substituted protein is then purified by conventional methods similar to the native protein and subjected to crystallisation trials. The MAD experiment is then performed at a synchrotron radiation source, as the absorption edge of selenium is at wavelengths generated by a synchrotron. Furthermore, to obtain meaningful MAD data, it is desirable to collect at three different wavelengths (to maximise anomalous differences) and this requires a tuneable X-ray source such as a synchrotron (Hendrickson WA, 1991).

3.6.1.5 Electron Density Maps, Refinement, Verification and Structural Analysis

An electron density map is a map showing the distribution of electrons. Since the electrons surround the nuclei of atoms it is possible for the atoms of the protein molecule to be fit into the electron density. This is known as model building and is usually performed on a computer graphics display. Once a round of model-building has been performed the model is then refined. The refinement software adjusts the model so that it best fits the data and ensures that the model is stereo-chemically correct (i.e. appropriate bond lengths and angles). Following this alternate cycles of model building and refinement are performed resulting in the final structure been determined which is then ready for structural analysis.

3.6.2 Cleavage of LES TCR acid/base zipper moiety

As mentioned beforehand, the levels of purity for the LES γδ TCR following Ni-NTA was insufficient for crystallisation trials. In addition, previous studies have highlighted that TCR crystallisation has not been successful with zipper chains attached to the corresponding TCR chains (Garboczi *et al*, 1992; Adams *et al*, 2005). It was highly likely that the presence of the flexible acid/base zippers would interfere with crystal packing and prevent crystal formation. This necessitated the removal of the acid/base zippers by cleaving with HRV 3C protease followed by additional purification before crystallisation screening trials could be attempted. The first stage in this process involved pooling together fractions eluted from the Ni-NTA column and buffer exchange into a buffer that was most suitable for the cleavage reaction, in this case 150 mM NaCl and 5mM Tris pH 8.

The ideal time point and concentration of 3C protease required to facilitate complete cleavage of the acid/base zippers from the corresponding individual chains were evaluated by carrying out replicate digests with differing protease concentrations. The resulting TCR digests were analysed by SDS-PAGE and the mobility of the TCR chains over different concentrations showed that the 1:50 dilution of protease enzyme was optimal for effective 3C protease cleavage activity (data not shown). The optimal digestion time was also calibrated by means of digesting the TCR protein and collecting analytical samples at differing time points which were then analysed by SDS-PAGE. Based on the gel, the optimal time for enzymatic digestion was determined to be two hours as protein mobility was unchanged after this time interval (Figure 3.14).

3.6.3 Crystallisation Trials of LES γδ TCR

Crystallisation trials of concentrated LES γδ TCR were carried out by using the commercial spare matrix kits and a mosquito nanolitre robot, with experiments performed at 23°C. The experimental method used was hanging drop vapour diffusion with the TCR concentration at 8.5-10 mg/ml. The purity of HRV 3C protease cleaved LES γδ TCR was evaluated by SDS-PAGE analysis after concentration and before crystallisation trials were undertaken (Figure 3.15). This approach led to the identification of two initial hits obtained with the PEG/ion (Figure 3.16) and Structure (Figure 3.17) screens. Unfortunately, these initial hits were difficult to reproduce due to the instability of LES γδ TCR in solution. Furthermore, continuous freeze thawing of LES γδ TCR sample exacerbated protein solubility resulting in loss of protein due to aggregation. Due to these problems, reproduction or optimisation of preliminary crystallisation hits was not possible.

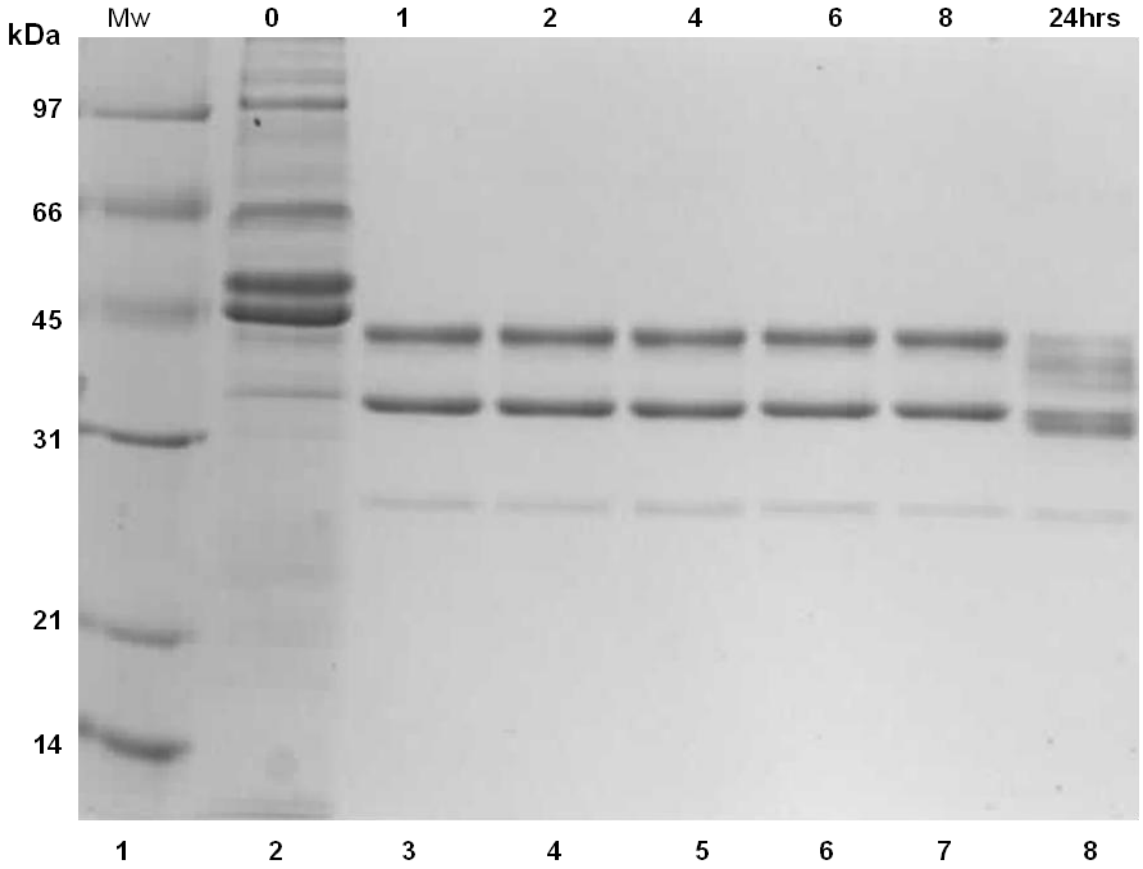


Figure 3.14 HRV 3C protease cleavage of LES γδ TCR sample analysed by SDS-PAGE. Time point analysis of 3C cleavage was performed over a period of 24 hours. Following protease cleavage with HRV 3C there is a significant downward shift of the TCR chains (Lanes 3-8) relative to undigested TCR (Lane 2), indicating removal of the acid/base zippers (6 kDa). After 24 hour incubation with the HRV 3C protease degradation of the TCR chains are observed (Lane 8).

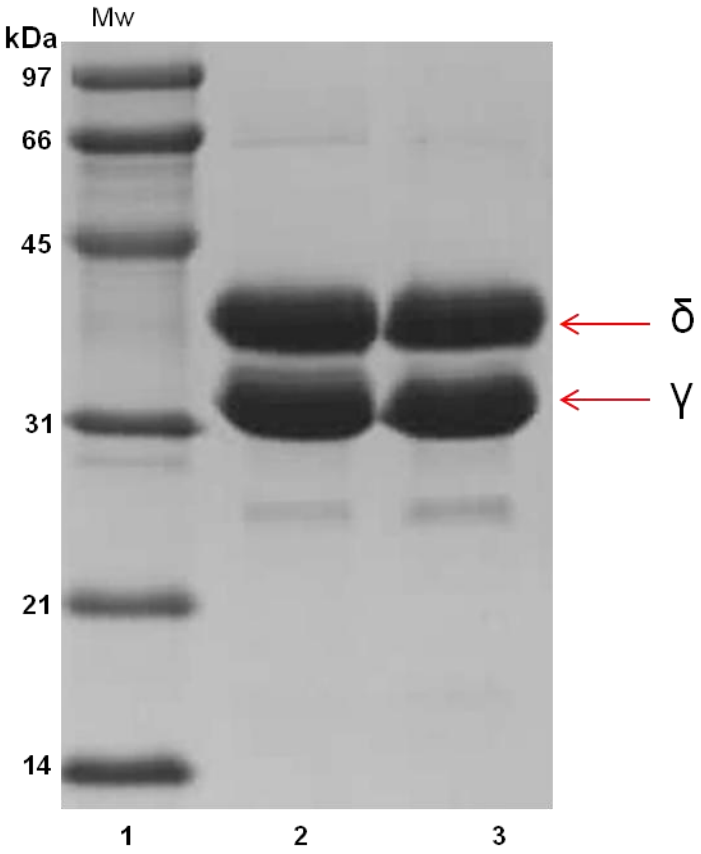


Figure 3.15 SDS-PAGE analysis of purified LES γδ TCR used for crystallization trials. Following cleavage with HRV 3C protease the acid/base zipper fragments were removed by Ni-NTA depletion and the LES γδ TCR purified by Ni-NTA beads. Lanes 2 and 3 represent concentrated forms of LES γδ TCR at 8.5 and 10 mg/ml, respectively. Molecular weight standards are shown in Lane 1, with masses indicated in kDa.

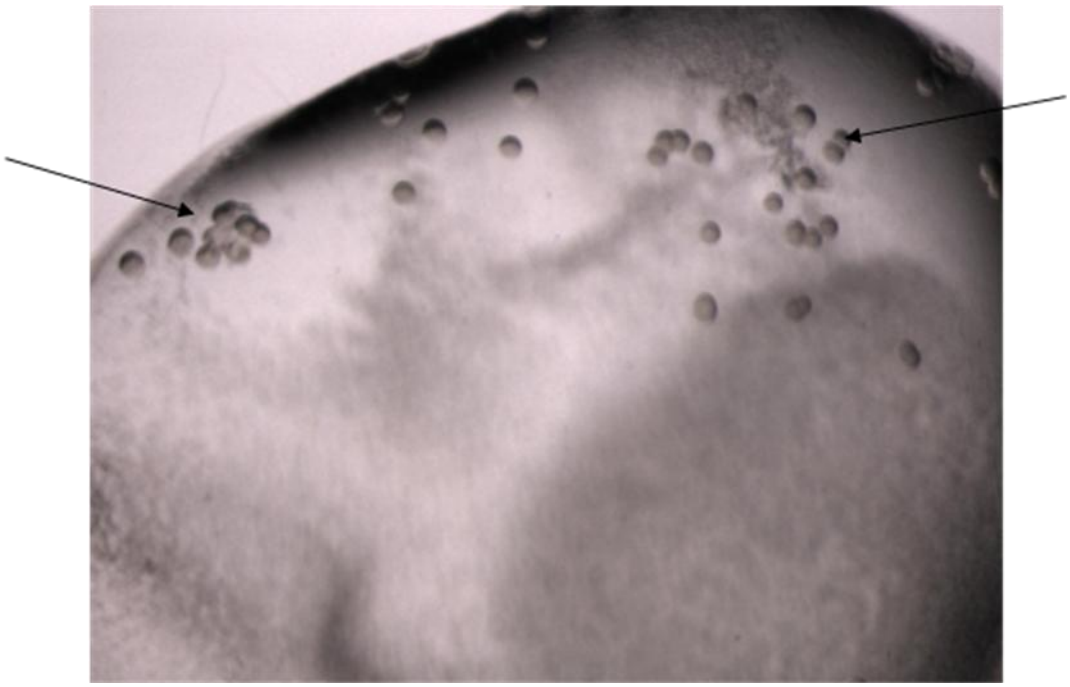


Figure 3.16 Generation of LES γδ TCR micro-crystals. Micro-crystals of LES γδ TCR (black arrows) were identified using the PEG Ion screen (Hampton Research). The drop was set up by mixing LES γδ TCR (8.5mg/ml in 150mM NaCl/5mM Tris pH8 buffer) with an equal volume of reservoir solution (condition number G4, 0.2M Sodium formate pH7, 20% PEG 3350). Micro-crystals were grown using the hanging drop vapour diffusion method and appeared after 14 days at room temperature.

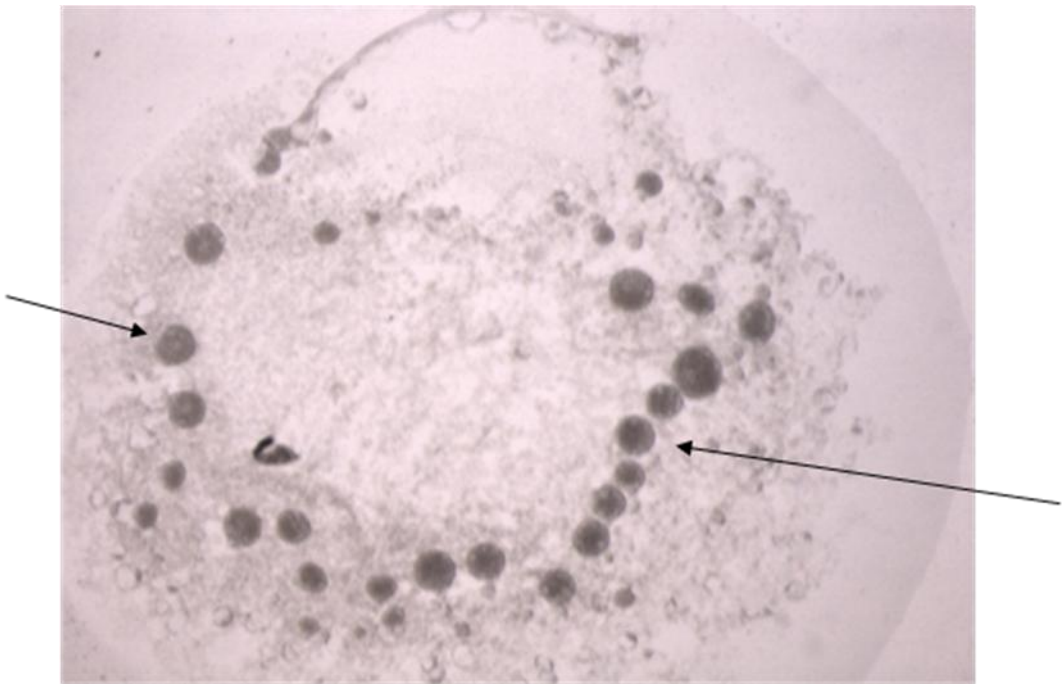


Figure 3.17 Generation of LES γδ TCR micro-crystals. Micro-crystals of LES γδ TCR (black arrows) were identified using Crystal Screen 2 screen (Hampton Research). The drop was set up by mixing LES γδ TCR (10mg/ml in 150mM NaCl/5mM Tris pH8 buffer) with an equal volume of reservoir solution (condition number H7; 0.2M Ammonium phosphate monobasic, 0.1M Tris pH8.5, 50% MPD). The micro-crystals were grown using the hanging drop vapour diffusion method and appeared after 14 days at room temperature.

3.7 Discussion

As touched on briefly earlier there were several reasons for using the prokaryotic expression system to generate LES and MAU γδ TCRs. Firstly, recombinant production using the *E. coli* expression system is relatively inexpensive relative to insect cell derived protein. Secondly, there are considerable differences in terms of the time-frame required to produce expressed proteins within the two expression systems. For example, recombinant expression with *Drosophila* can take approximately several weeks, whereas protein expression within the *E. coli* system can be achieved within three days. Finally, the protein yields from *E. coli* far exceed those than generated by *Drosophila*. Despite the obvious advantageous properties conferred the *E. coli* system, there was a requirement to refold the LES and MAU γδ TCRs back to their native conformation. This is partly due to the fact that the prokaryotic expression machinery is not fully equipped (e.g. lacking key post translational modification enzymes and chaperones) to refold complex human proteins and as such the TCRs were expressed as insoluble inclusion body aggregates. Indeed, previous studies have highlighted several strategies that have successfully been employed to correctly refold αβ (Simpson *et al*, 2011) and γδ (Garboczi *et al*, 1996) TCRs back to their native heterodimer conformation in yields sufficient for structural and binding applications.

In this study use of similar approaches provided limited success in refolding LES and MAU γδ TCR. Although, it was possible to generate significant levels of LES and MAU TCR inclusion body material (typically 100 mgs of each chain were produced per litre of LB culture), the yields following purification by anion exchange chromatography were extremely low for downstream functional and structural studies. One possible reason for low recovery yields was attributed to the fact that the initial inclusion body protein material, particularly for

the LES TCR γ and δ chains, was impure which may have interfered with the renaturation process. One solution for circumventing such a problem is to further purify the inclusion bodies, but past experiences have led to the conclusion that although such strategies can provide pure protein, the recovery yields are extremely low. Furthermore, the variation in protein purity levels between inclusion body preparations may have exacerbated the inefficient refolding of the γδ TCR. Another major problem associated with the generation of TCR using the prokaryotic system is the lack of TCR stoichiometry observed following refolding. Indeed, SDS-PAGE analysis of the γδ TCR refolded by *in vitro* dilution methods frequently revealed higher levels for the δ chain relative to the γ chain, raising considerable doubts as to whether the recombinantly derived γδ TCR is conformationally functional. One possible strategy for achieving correct stoichiometry of the TCR is supplementing the refold buffer with excess quantities of the γ chain relative to the δ chain. This may then provide sufficient levels of refolded γ chain in the refold mixture for pairing with the δ chain and facilitate stoichiometric association of the two chains. Collectively, these observations suggested that use of the prokaryotic expression system to generate recombinant γδ TCRs for structural and functional studies was impractical given the time constraints of this study.

Due to the huge number of technical obstacles encountered in generating large amounts of heterodimeric γδ TCR protein with the prokaryotic system, an alternative approach with the *Drosophila* S2 cells was utilised. Indeed, the *Drosophila* expression system provided 2 mg of LES and MAU heterodimer TCR. Despite the lower yields subsequent dot blot analysis confirmed correct pairing of each TCRs by using the 2H11 antibody that specifically bound to the acid/base zipper moieties. The fractions eluted following Ni-NTA purification proved to be highly promising with two prominent bands observed, corresponding to each chain. There

was however several larger molecular weight bands present suggesting that the Ni-NTA purification protocol was sub-optimal. In the future it may be possible to improve the recovery yields by modifying the wash step to more efficiently remove non-specific proteins bound to the Ni beads. For instance, by washing with PBS buffer containing 20 mM imidazole (rather than 10 mM) there is an increased likelihood of generating purer TCR reagents for the subsequent cleavage step with 3C protease. Alternatively, if this proves to be unsuccessful it may be possible to remove the contaminants by including an additional analytical procedure following Ni-NTA chromatography such as size exclusion chromatography.

Nevertheless, despite these issues it was possible to generate sufficient amounts of 3C cleaved LES γδ TCR for preliminary crystallisation trials. Although the screening process yielded several crystallisation hits, they were deemed very poor suggesting that considerable optimisation was required in order to produce diffraction grade crystals. Optimizing all of the most obvious parameters which included pH, salt and protein and precipitant concentration continued to produce crystals displaying poor size and morphology. A review of the procedures used before and after crystallization trials has highlighted several issues, that when addressed, may lead to enhanced crystal growth. Firstly, the stability of the LES γδ TCR was relatively poor in that it displayed an increased propensity to aggregate following multiple freeze thaw cycles, particularly at high concentrations, which was necessary for promoting nucleation. These stability issues most likely stemmed from the choice of the storage buffer used for the LES γδ TCR. Indeed, recent studies have demonstrated that incorporating thermal melting assays into the crystallization screening experiments enhanced the number of initial crystallization leads (Ericsson *et al*, 2006). A similar approach could be applied to LES γδ

TCR to identify a more stable storage buffer suitable for crystallization experiments. Secondly, future crystallisation trials could encompass using other physical techniques including dialysis and batch. Although both methods can induce the protein solution to achieve super-saturation, the underlying principles of these techniques vary. The batch method crystallisation technique consists of mixing concentrated protein solution with crystallisation reagent in a closed vessel to produce a final supersaturated concentration, which may ultimately lead to crystallisation. In contrast, dialysis techniques exploit diffusion and equilibration of small precipitant molecules through a semi-permeable membrane as a manner of gradually approaching the desired concentration at which protein crystallizes. Thirdly, it is clearly evident that the *Drosophila* derived LES γδ TCR is heavily glycosylated as there were significant shifts in migration observed for both chains relative to their sequence derived molecular weight. The glycosylation associated phenomenon observed for LES γδ TCR is unsurprising given that there are 2 γ chain and 3 δ chain predicted N-linked glycosylation sites for the gamma and delta chain, respectively. Indeed, it is a well established that the chemical and conformational heterogeneity of the carbohydrate moieties can inhibit protein crystallisation, thereby providing a rationale for intransigent crystallisation of the LES γδ TCR. However, there are several approaches that could be employed to overcome the glycosylation problem and enhance crystallisation of the LES γδ TCR. One possibility is the use of PNGase F or Endonuclease H to completely deglycosylate the LES γδ TCR sample prior to the crystallisation trials.

A second approach involves expressing the LES γδ TCR transiently in insect cells with *N*-glycosylation processing inhibitors such as tunicamycin. This approach has been successful in facilitating the crystallisation and structural determination of G8 γδ TCR in complex with

T22, which was treated with tunicamycin (Adams *et al*, 2005). This may allow the TCR to fold correctly, but more importantly the protein is likely to be increasingly sensitive to enzymes, such as endoglycosidase H, that reduce the *N*-glycans to single uniform residues, and thus enhance crystallization. Strikingly, these and similar strategies have been employed in a number of structural biology consortium initiatives (Chang *et al*, 2007; Adams *et al*, 2005). Finally, proteins that are refractory to crystallization are more likely to crystallise in the presence of their ligands as it allows the protein to adopt a more stable form in solution (Warke and Momany, 2007). Indeed, the validity of co-crystallization has been highlighted by other members within the Willcox laboratory. For example, it has been possible to circumnavigate crystallisation of intransigent HLA-A2-peptide complexes with the incorporation of its natural ligand, LILBR1 (personal communication Dr Fiyaz Mohammed). With this in mind, attempting the co-crystallisation of the LES γδ TCR with its cognate ligand, EPCR (more details in chapter 4), could stabilise the structure in solution and enhance crystallisation.

CHAPTER 4 –

MOLECULAR CHARACTERISATION OF A HUMAN γδ T CELL RECEPTOR LIGAND

Chapter 4: Molecular characterisation of a human γδ T cell receptor ligand

4.1 Introduction

As described in Chapter 3, the Déchanet-Merville group generated a panel of $V\delta2^{neg}$ γδ T cell clones from immunosuppressed solid organ transplant patients with acute CMV-infection and determined their reactivity to CMV-infected cells and tumour cell lines, with the goal of identifying the ligands recognised by this population of γδ T cells. At present, there are several strategies that could be used to identify potential TCR ligands, including transfection of non-target cell lines with a cDNA library to induce expression of the TCR ligand; transfection of target cell lines with an siRNA library to inhibit expression of the TCR ligand; or generation of blocking antibodies which would bind the TCR ligand and inhibit T cell activation. Each strategy has advantages and disadvantages. One such strategy adopted by Dr Déchanet-Merville's laboratory was to attempt to generate blocking antibodies. Antibodies were generated in mice immunised with the HT29 human colorectal carcinoma cell line, which are recognised effectively by the LES γδ clone and LES JRT3 transfectants. One antibody blocked activation of the LES γδ clone by HT29 and other target cell lines as well as by CMV-infected fibroblasts, and was designated 2E9 (Willcox *et al*, 2012, Appendix I). Crucially, 2E9 blocking was TCR-specific, since it did not block recognition of HT29 by the MAU γδ T cell clone, suggesting 2E9 was not blocking a common costimulatory or NKR ligand required for activation of both γδ clones, and instead that 2E9 was highly likely to be blocking the ligand of the LES γδ TCR. Interestingly, whereas the 2E9 ligand (2E9L) was essential for activation of LES γδ T cells, it did not appear to be sufficient, as some $2E9L^{+}$ cell lines were not productively recognised by the LES clone (Table 3.1). This work highlighted 2E9 as a potentially key reagent that may allow identification of the LES γδ TCR ligand.

Subsequent immunoprecipitation experiments were performed by Dr Carrie Willcox to identify 2E9L. $2E9^{+}$ cell lines were surface biotinylated and lysed, and a protein of 37kDa was specifically immunoprecipitated from HeLa cells by 2E9. The band was isolated and upon analysis by Fourier transform-ion cyclotron resonance mass spectrometry, four peptides derived from endothelial protein C receptor (EPCR) were identified (Willcox *et al*, 2012). EPCR was confirmed as the 2E9L in several ways. A commercial anti-EPCR antibody was used to western blot the band specifically immunoprecipitated by 2E9 in HT29 and HeLa cells. In addition, a commercial anti-EPCR polyclonal antibody blocked staining of target cells by 2E9, and also blocked activation of the LES clone by HT29 target cells to a similar degree as 2E9 (see Willcox *et al*, 2012 Appendix I Figures 2c-e). Finally, BIAcore analyses proved that 2E9 specifically bound immobilised EPCR protein with high affinity.

While it was apparent that EPCR was the 2E9L, it had not been confirmed as the ligand for the LES γδ TCR. To address this, the cDNA for the ectodomain of EPCR was cloned into the pMTBiPHis vector and expressed in the *Drosophila* expression system (DES). Soluble EPCR (sEPCR) was purified (Appendix I, Supplementary Figure 1a) and tested in SPR studies for binding to the *Drosophila* derived LES γδ TCR (described in Chapter 3). Greater binding responses were observed following injection of sEPCR over immobilised LES γδ TCR than over immobilised MAU γδ TCR or an irrelevant αβ TCR control (Appendix I, Figure 3). This confirmed EPCR is a direct ligand for the LES γδ TCR. Equilibrium binding analyses revealed an apparent dissociation constant (Kd) of 87μM, which is low but within of the range demonstrated for αβ TCR-pMHC interactions (1-100 μM) (Rudolph *et al*, 2006). Notably, EPCR produced in human 293T cells also specifically bound LES γδ TCR with a similar affinity.

EPCR is a well established component of the protein C anti-coagulation pathway. Protein C binds to EPCR on endothelial cells and is cleaved by the thrombin-thrombomodulin complex to form activated protein C (APC). APC then inactivates factor V in the coagulation cascade, which ultimately results in a decrease in fibrin formation and clotting (Esmon, 2010) (Figure 4.1). Moreover, EPCR is also required for the anti-inflammatory and endothelial barrier protective role of APC.

EPCR is a type I transmembrane protein with approximately 20% sequence homology to CD1d. Despite this low sequence homology, the structure of EPCR (Oganesyan *et al*, 2002) highlights a high degree of structural similarity to CD1d (Koch *et al*, 2005) (Figures 4.2a-b): EPCR consists of α1 and α2 helices, positioned above a β-sheet platform, and between the two alpha helices is a hydrophobic channel enabling binding of a lipid. In the crystal structure, EPCR was shown to bind either Phosphatidylethanolamine (PE) or Phosphatidylcholine (PC). However, in contrast with CD1d, EPCR lacks an α3 domain and is not associated with β_2-microglobulin (β_2M). In addition, the lack of a β_2M domain conceivably enables the α1/α2 domain to be more flexible with respect to the membrane, which may facilitate binding to APC (Figures 4.3a-b).

Although EPCR has not previously been recognised as having an immunological function, it fulfils the hypothesis that ligands for $V\delta2^{neg}$ cells would be present in both HCMV infection

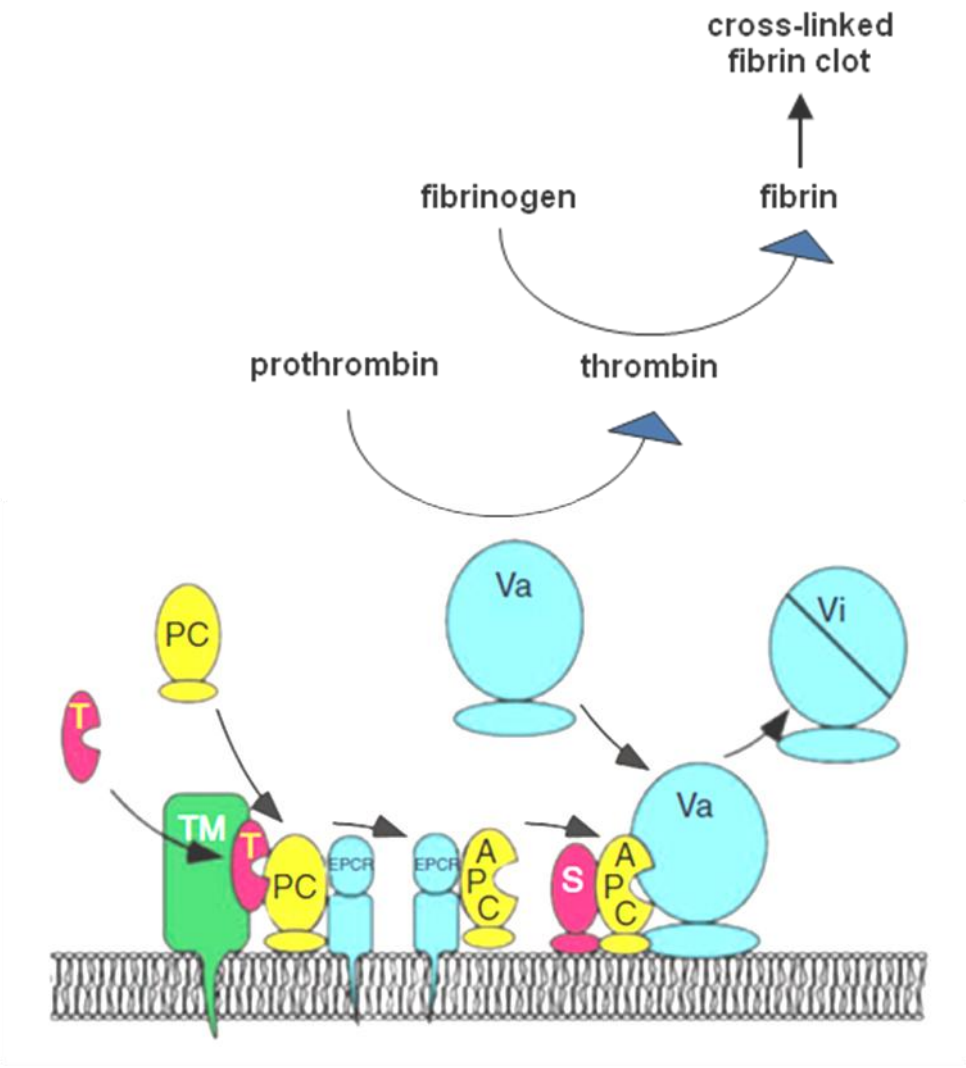


Figure 4.1 Cartoon representation of EPCR-dependent protein C (PC) activation by the thrombin (T) – thrombomodulin (TM) complex. Activated protein C (APC) binds to protein S (S) and inactivates factor Va, thereby down regulating the downstream coagulation cascade (adapted from Esmon, 2010).

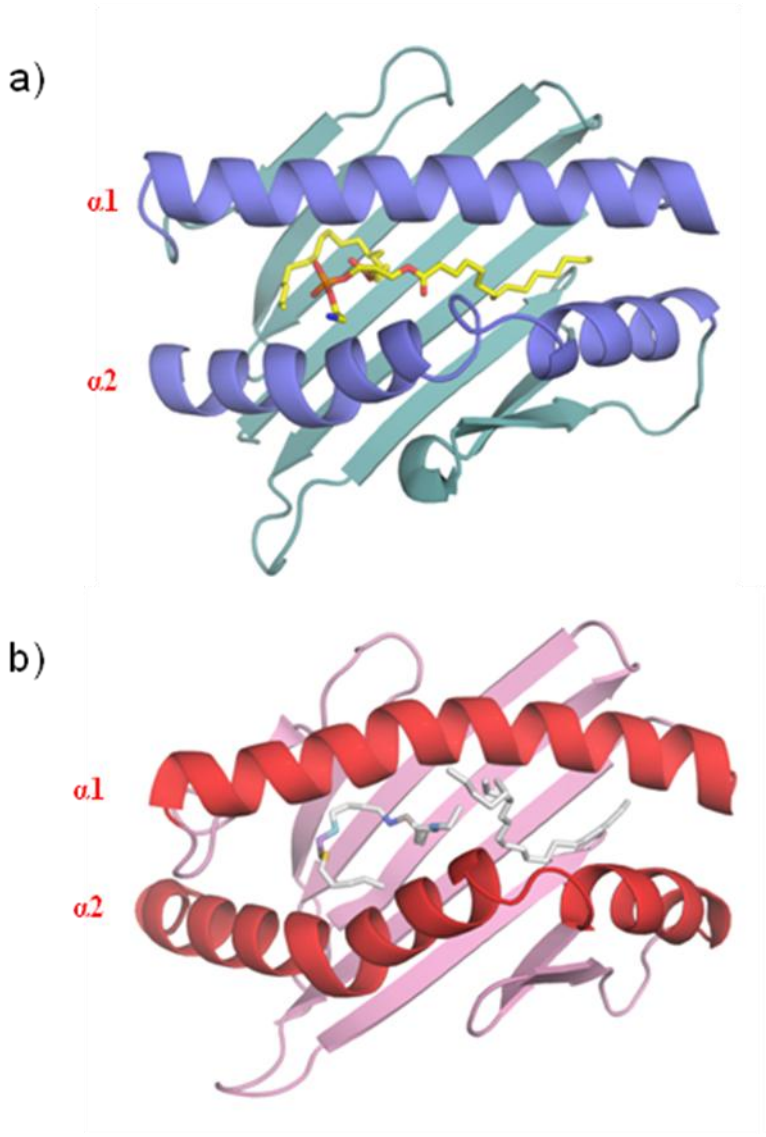


Figure 4.2 Structural comparison of the lipid binding platform of a) EPCR (Oganesyan *et al*, 2002) and b) the α1/α2 domain of CD1d (Koch *et al*, 2005). In each case the bound lipid within the hydrophobic groove is highlighted in stick format.

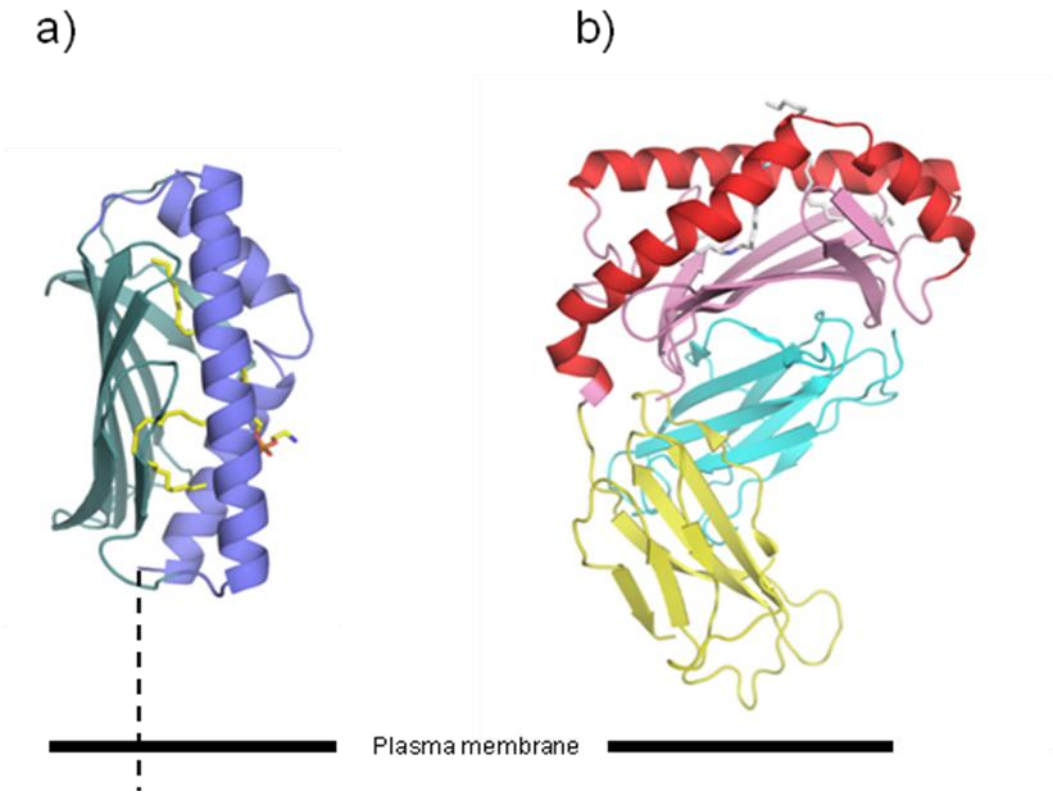


Figure 4.3 Structural comparison of a) EPCR (Oganesyan *et al*, 2002) and b) CD1d (Koch *et al*, 2005) as each molecule is proposed to be presented on the plasma membrane (side view). The α1/α2 platform of CD1d is held parallel to the plasma membrane because of interactions between the underside of the b-sheet and the α3 domain/β_2M. EPCR is thought to be anchored to the membrane by a flexible stalk, allowing it to adopt different conformations, including the one shown in which the α1/α2 platform is perpendicular to the plasma membrane. In each case the bound lipid within the hydrophobic groove is highlighted in stick format.

and cancer; firstly, endothelial cells are a key target for HCMV infection and conceivably expression of EPCR on the endothelium might suggest surveillance of this site by γδ T cells. In addition, EPCR has been shown to be over-expressed on cancer cell lines and in tissue sections of some human cancers (Scheffer *et al*, 2002), suggesting recognition of EPCR by γδ T cells as a possible mechanism of tumour immunosurveillance.

Despite the fact that EPCR is closely related to the CD1 family that present lipid antigens to T cells, the mechanism by which EPCR is recognised by LES γδ T cells remains unclear and has important implications. It is intriguing that some cell lines stain with 2E9 but do not activate the LES γδ clone or JRT-3 transfectant, suggesting that expression of EPCR is necessary but not sufficient for recognition by LES T cells (Table 4.1). This could be explained by two different mechanisms (Figure 4.4). Firstly, activation by $EPCR^+$ cell lines could be dependent on EPCR-intrinsic factors, such that EPCR may be expressed in a different and more stimulatory form on the surface of activatory cells relative to non-activatory $EPCR^+$ cells. Perhaps most likely, the homology between EPCR and CD1d suggests that EPCR may have the ability to present a stress lipid to LES γδ TCR (Figure 4.4a). A putative stress lipid might only be present on activatory cancer lines (versus non-activatory $EPCR^+$ lines) and might conceivably be upregulated after CMV infection of target cells. Alternative discriminatory factors could include differential EPCR glycosylation, which could directly affect binding to the TCR; or the distribution of EPCR on the plasma membrane, or its association with other proteins on the cell surface, might influence recognition of target cells by LES γδ T cells. In contrast, EPCR-extrinsic factors could instead determine whether $EPCR^+$ cell lines can activate LES γδ T cells (Figure 4.4b). In this scenario, EPCR expressed on stimulatory and non-stimulatory $EPCR^+$ lines would be identical, however EPCR-extrinsic factors such as differential expression of costimulatory

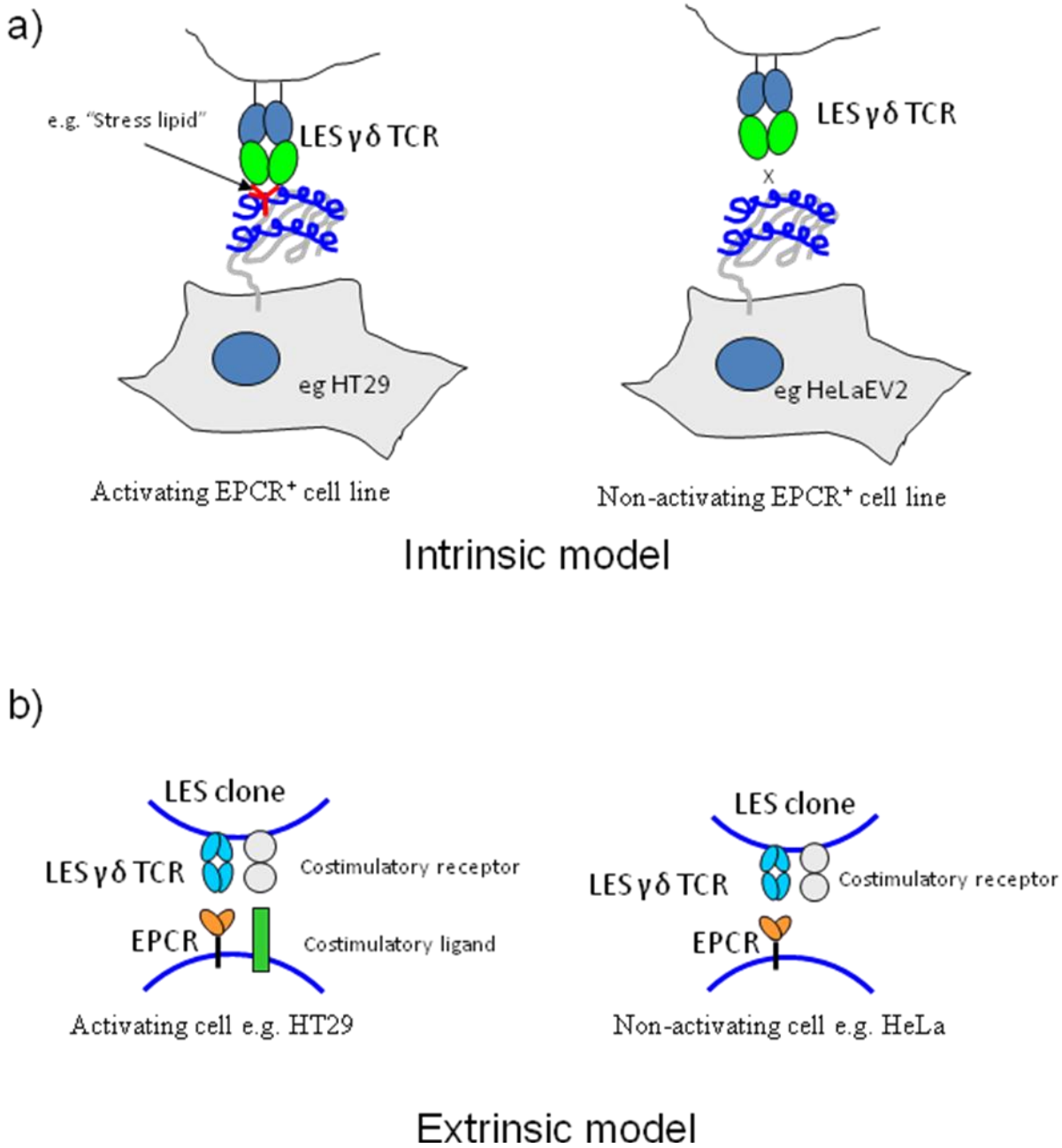


Figure 4.4 Models of LES activation by EPCR expressing tumour cells. EPCR is necessary but not sufficient for LES activation. The difference in response of the LES clone to activating and non-activating target cell lines could be explained by EPCR-intrinsic factors (a) or EPCR-extrinsic factors (b). EPCR-intrinsic factors could include presentation of a stress lipid which is only expressed by activating cell lines. EPCR-extrinsic factors could indicate the requirement for additional costimulatory ligands, present on $EPCR^+$ activating cells, but not expressed on $EPCR^+$ non-activating cells.

ligands could render stimulatory EPCR^{+} cell lines more conducive to productive recognition by LES T cells.

Cell Line	LES	2E9	EPCR expression
HT29	✓	✓	+
K562	✓	✓	+
U937	✓	✓	+
HUTU80	X	X	-
HUH7	X	X	-
U373	X	✓	-
293T	X	✓	-
HeLa EV2	X	X	-
HeLa PF	X	✓	+
SKMES	X	✓	+
MRC5	X	✓	

Table 4.1 Staining of cell lines with 2E9 mAb and recognition by LES γδ TCR. *List of Cell lines which were analysed for recognition by the LES γδ clone or JRT-3 LES transfectants. The presence of EPCR on the cell line was also determined and recognition by 2E9 mAb.*

Arguably the most compelling of these explanations is that EPCR acts as a presentation element, enabling LES T cells to recognise "stress lipids" selectively present on stimulatory target cell lines. Importantly, recognition of a stress lipid in EPCR would require TCR interaction with the solvent exposed surface of the α1/α2 helices in a fashion analogous to that of TCR/CD1 interaction, whereby bound lipids are discriminated directly by the TCR, based on their lipid head group (Girardi and Zajonc, 2012). In contrast, a TCR binding site on EPCR that was distinct from this face of the α1/α2 platform would effectively exclude direct recognition of EPCR-bound lipids. This would suggest that instead of serving as a restriction element for stress-regulated lipids, EPCR may act as a stress ligand itself. It would also highlight the potential importance of EPCR extrinsic factors such as expression of

costimulatory molecules as playing a critical role in activation. Therefore, in order to shed light on these issues, the key aim of this chapter was to define the binding mode of the LES Vγ4Vδ5 TCR for EPCR. The primary approach adopted was based on direct protein-protein binding experiments using SPR, in combination with alanine scanning mutagenesis of the EPCR molecule. A second, parallel and complementary strategy was to attempt crystallisation of the LES γδ TCR in complex with EPCR.

4.2 Soluble EPCR expression

The expression of soluble EPCR was developed in the DES, the expression system that was initially used to establish the LES TCR directly bound EPCR. In addition, by allowing endogenous *Drosophila* lipids to be loaded into the molecule, it would therefore circumvent the difficulty of artificially loading a lipid into *E. coli*-produced, refolded EPCR. Indeed, previous experimental work has shown the intrinsic difficulty of loading lipids into artificially refolded protein molecules such as CD1 (M. Salim unpublished observation). The cDNA for the ectodomain of EPCR was cloned into the pMTBiP vector containing a BirA and His tag by Dr Carrie Willcox. Soluble EPCR was expressed in S2 cells, and then purified by nickel chelate chromatography and biotinylated using the BirA enzyme. The purified EPCR was shown to bind both the 2E9 anti-EPCR mAb, as well APC, the known ligand for EPCR (Appendix I Supplementary Figures 1a-c).

It has already been demonstrated that sEPCR binds immobilised LES γδ TCR (Appendix I Figure 3). To confirm this interaction in the reverse orientation soluble TCR was injected over a flow cell immobilised with EPCR. As a negative control, biotinylated HLA-A2 was

immobilised on a control flow cell. The resulting sensorgrams revealed considerably higher responses after injection of soluble LES γδ TCR over immobilised EPCR than over HLA-A2, indicating specific binding. Representative binding curves and Scatchard plots are shown in Figure 4.5 and multiple equilibrium binding analyses gave an apparent Kd of 91.3μM +/-46.7 μM (n=30), consistent with the value originally obtained in the reverse orientation (87 μM) (Appendix I). To determine whether LES γδ TCR recognised a conformational epitope on EPCR, soluble biotinylated EPCR was denatured. Specific binding responses were observed to native EPCR but not EPCR pre-treated with DTT at 60°C (Figure 4.6a). To confirm the EPCR was completely denatured, 2E9 mAb was injected and specifically interacted with native EPCR but binding to denatured EPCR was greatly reduced (by approximately 70%, Figure 4.6b). Therefore this confirmed that LES γδ TCR binds a conformational epitope on EPCR.

EPCR has four predicted N-linked glycosylation sites (Asn-X-Ser/Thr: N30, N47, N119 and N155), and differences in the extent of EPCR glycosylation have been observed in activating and non-activating target cell lines by SDS-PAGE and Western blotting (Appendix I Figure 2), although no obvious correlation between glycosylation state and activity emerged. However, it was important to determine whether the glycosylation status of EPCR was critical for the binding of LES γδ TCR, particularly as insect cell expressed proteins lack complex-type N-glycans present in mammalian proteins and instead consist mainly of high mannose structures, and therefore conceivably altered glycosylation status could be an important component of recognition. Purified EPCR was treated with peptide N-glycosidase F (PNGase F), which cleaves carbohydrate chains from the asparagine residue. Treatment of both native and reduced EPCR with PNGase F resulted in an increased mobility of the

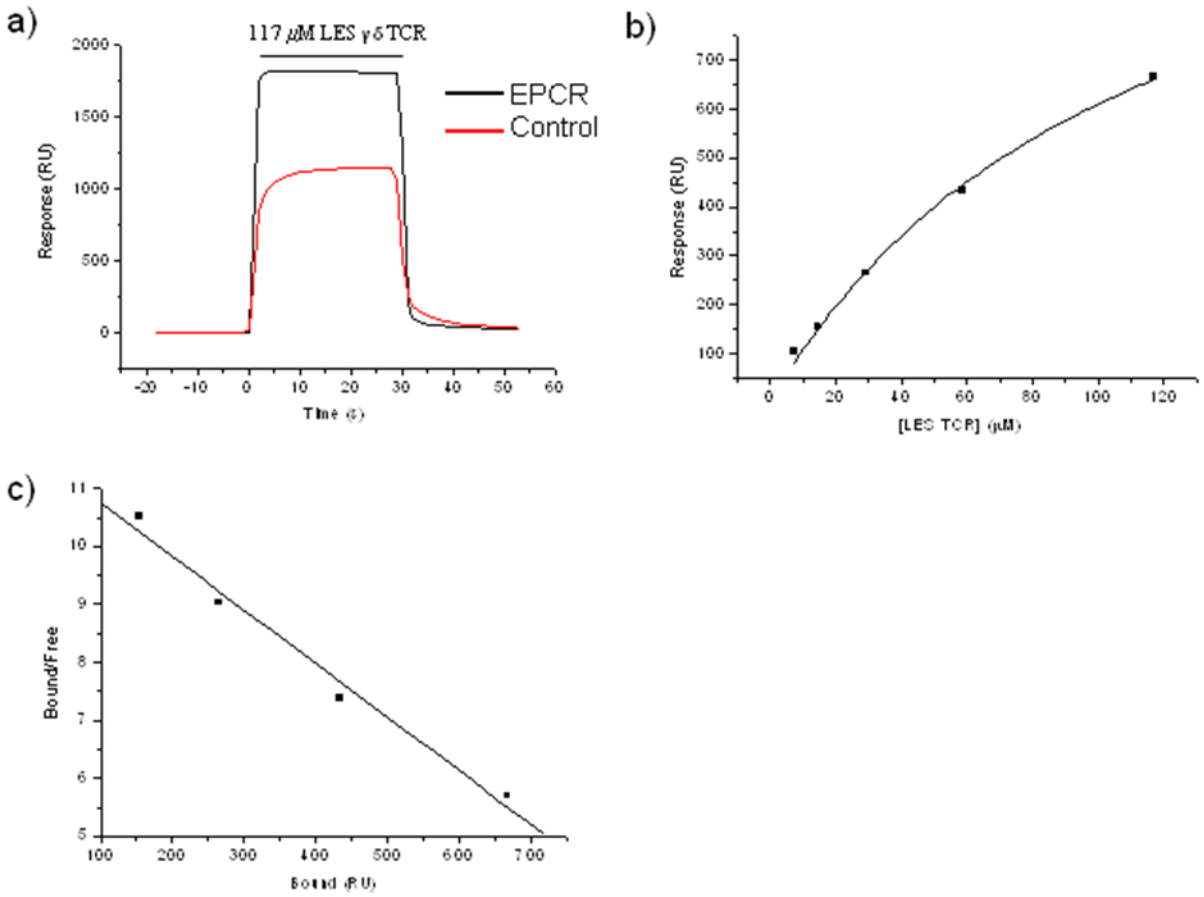


Figure 4.5 Soluble LES γδ TCR binds EPCR in the reverse orientation. a) Representative binding trace of LES γδ TCR (117 mM) injected over immobilised EPCR (1774 RU) or control HLA-A2 (1804 RU). b) Equilibrium binding analysis, involving injection of a range of LES γδ TCR concentrations (7.3-117 mM), allowed the calculation of an apparent dissociation constant (Kd) of 112.1 mM in this experiment. c) A Scatchard plot of the same data gave a Kd of 107.5 mM. This experiment was repeated (n=30), giving an average Kd of 91.3 +/- 46.7 mM.

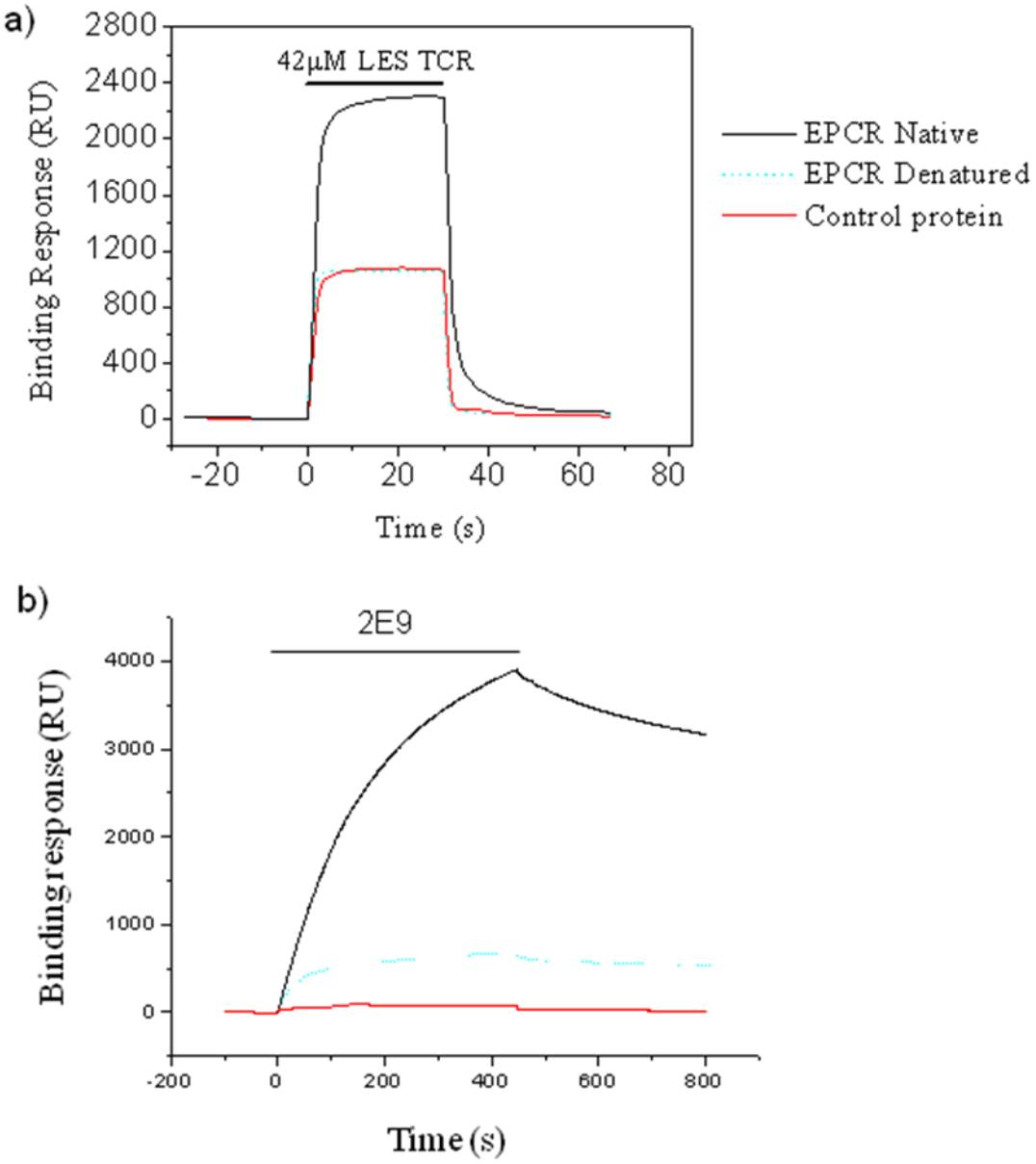


Figure 4.6 LES γδ TCR binds to EPCR in a conformational manner. a) LES γδ TCR was injected over immobilised native EPCR (2405 RU), denatured EPCR (1911 RU) and HLA-A2 control (2239 RU). To denature EPCR, purified EPCR was treated with 5mM DTT at 60°C for 1 hour then treated with 20mM iodoacetamide for 30 minutes in the dark to prevent disulphide bonds reforming. b) 2E9 (2 mM) was injected over immobilised native and denatured EPCR and control HLA-A2.

protein on SDS-PAGE (the major speciesmigrated at ~25 kDa, the predicted molecular mass of the EPCR protein backbone) (Figure 4.7a), suggesting the majority of carbohydrates were removed under both reducing and non-reducing conditions. Notably, PNGase F treatment of native EPCR highlighted that the glycosylation of EPCR did not play a role in LES γδ TCR binding, as glycosylated and deglycosylated EPCR bound TCR with similar affinity (Figures 4.7b-c). Also, PNGase F treatment of EPCR did not affect overall conformation of the structure as 2E9 bound deglycosylated EPCR equivalently to native EPCR (Figure 4.7d).

Based on these results, it is clear that the LES γδ TCR binds a conformation-dependent protein epitope on EPCR. Therefore, to further probe the binding mode of LES γδ TCR and EPCR, and examine whether lipid presentation was critical for TCR recognition, mutagenesis of the α1 and α2 helices of EPCR was undertaken. The amino acids selected for alanine scanning mutagenesis were solvent exposed, and hence unlikely to be important in stabilising the EPCR structure, and more likely to interact with the TCR. This selection process was made possible by detailed examination of the EPCR crystal structure (Oganeysan *et al*, 2002) (Figure 4.2). Using this strategy 17 initial EPCR mutant constructs were designed on the α1 (Figure 4.8a) and α2 (Figure 4.8b) helix (Table 4.2). To gain a more comprehensive profile of the LES γδ TCR/EPCR mode of binding, 8 further mutants were generated, in which additional regions of the EPCR were examined including the N-terminus and the β-sheet face opposite the antigen binding platform (Table 4.2; Figure 4.9).

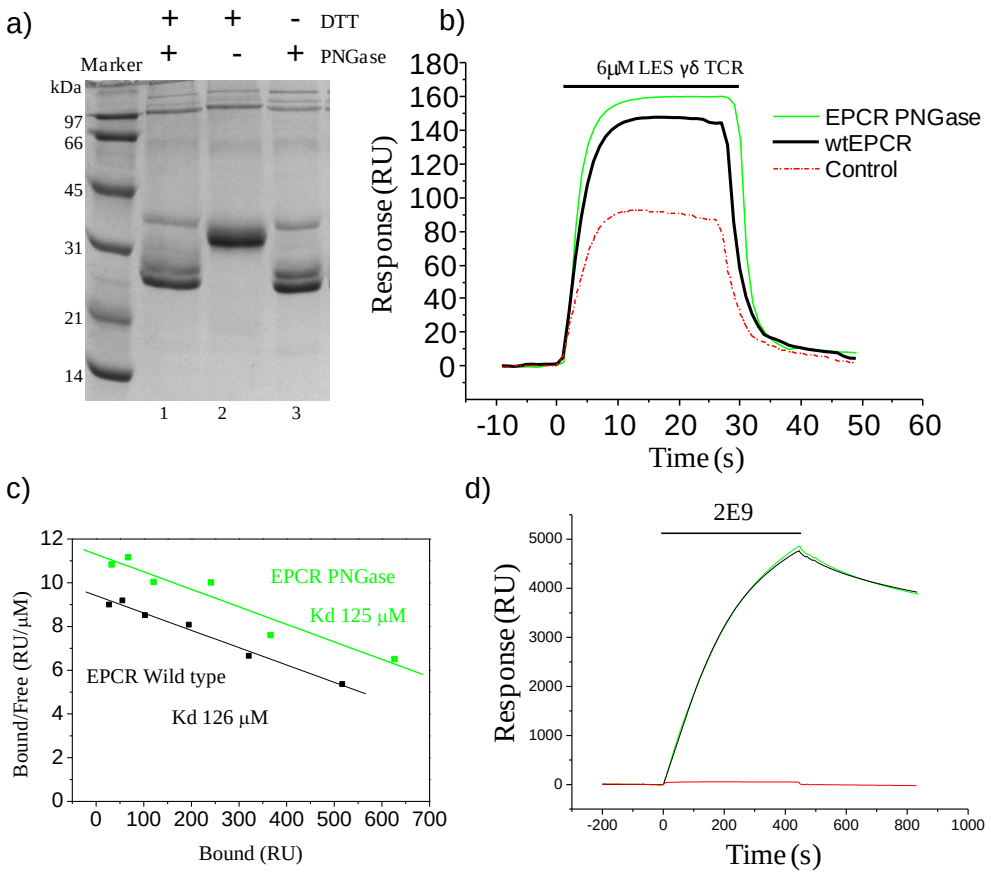


Figure 4.7 a) SDS-PAGE analysis of EPCR after PNGase F treatment. Lane 1 and 2 show DTT-reduced EPCR treated with and without PNGase F, respectively; lane 3 shows PNGase F-treated native EPCR; a shift in mobility is observed after removal of carbohydrates from the protein under both native and reducing conditions. b) SPR profile of LES γδ TCR injected over immobilised native EPCR (3300 RU), PNGase F-treated EPCR (3700 RU) and HLA-A2 control (3100 RU). c) Scatchard plots were generated following injection of LES γδ TCR over a range of concentrations (3-96 mM) and the Kd for wtEPCR (126.3μM) and PNGaseF treated EPCR (125 mM) were calculated d) 2E9 mAb (2 mM) was injected over native and denatured EPCR, and control HLA-A2.

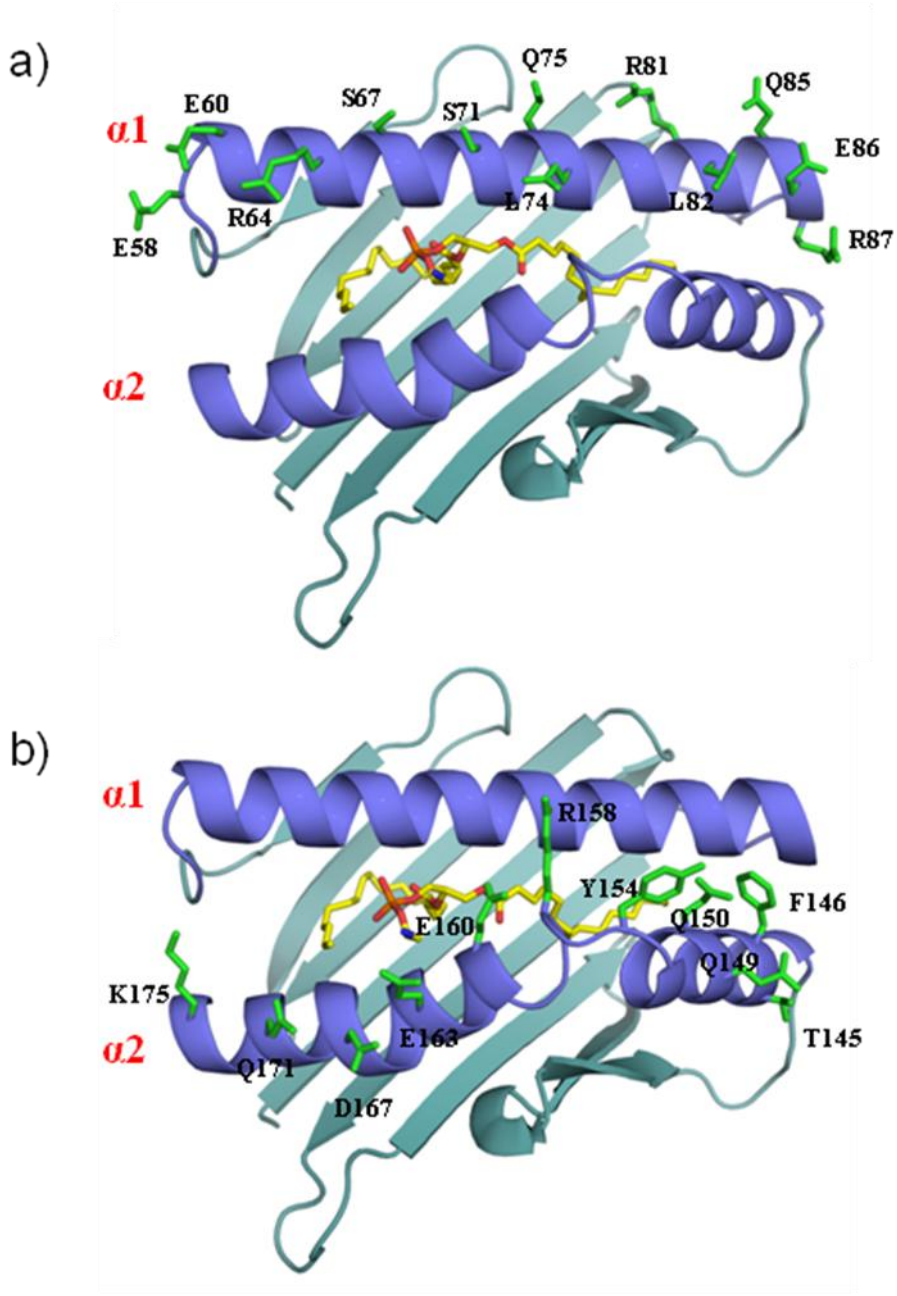


Figure 4.8 Lipid binding platform of EPCR, highlighting solvent exposed amino acids. Highlighted residues (side chains are shown in green in stick format) on the a1 helix (a) and a2 helix (b) were selected for alanine scanning mutagenesis to examine their effect on LES γδ TCR binding. Lipid is highlighted in stick format (yellow).

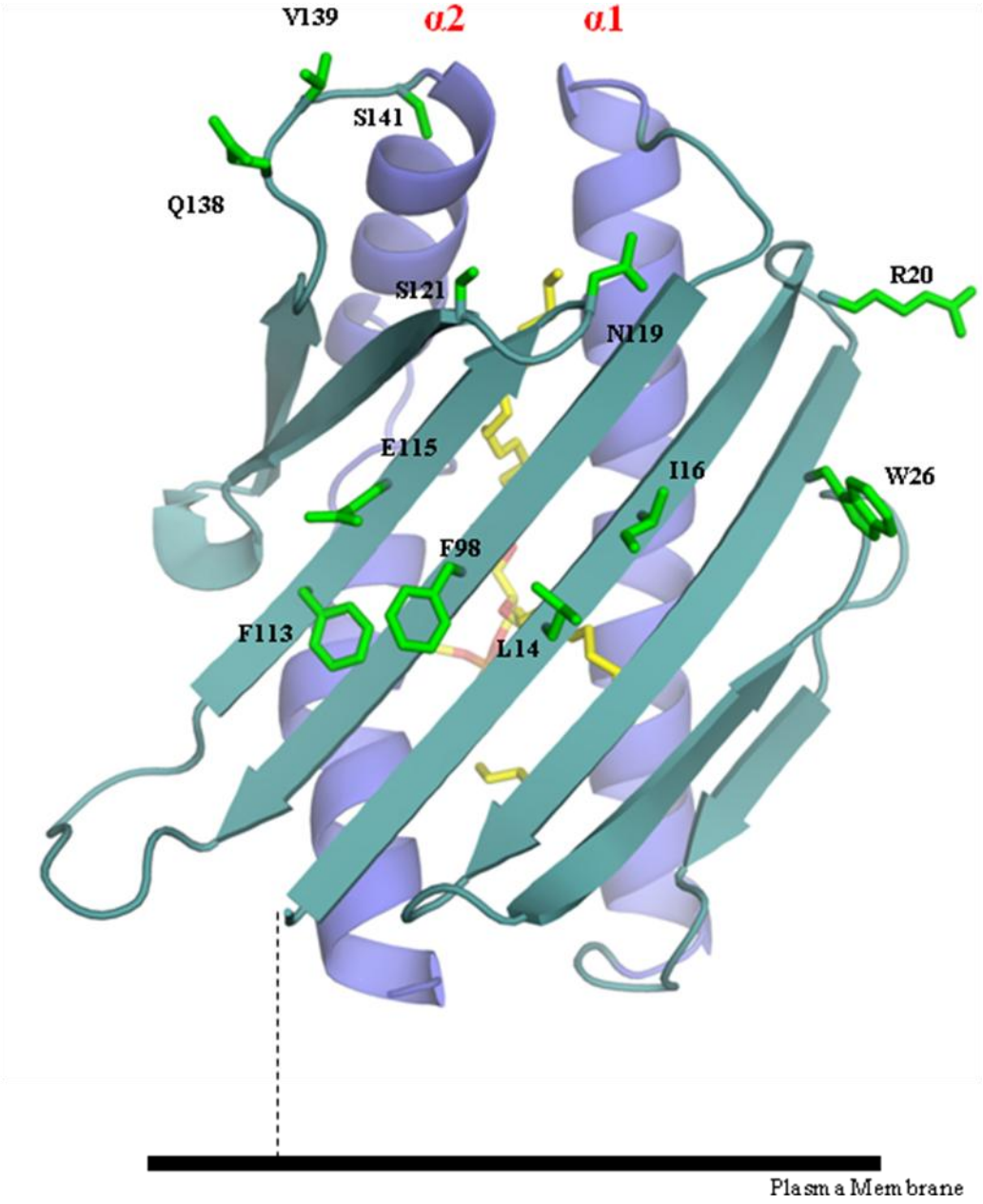


Figure 4.9 Underside view of the EPCR molecule highlighting solvent exposed residues on the β-sheet and interstrand loops that were targeted for alanine scanning mutagenesis. Selected residues are depicted as stick figures in green. Highlighted in yellow is the lipid bound to EPCR. Dashed line represents the stalk region of EPCR.

α1 helix	α2 helix	β-sheet
E58A + E60A	T145A	L14A + I16A
R64A	F146A	R20A
S67A	Q149A	W26A
S71A + L74A	Q150A	F98A
Q75A	Y154A	F113A + E115A
R81A+ L82A	R156A	N119A
Q85A	E160A	S121A
E86A + R87A	E163A	Q138A + V139A + S141A
	D167A + Q171A + K175A	

Table 4.2 Solvent exposed amino acids mutated on EPCR. Listed are the residues selected for alanine scanning mutagenesis. *Exposed residues located on the α1 and α2 helices were selected, as they are not predicted to stabilise the EPCR molecule and could be in contact with the LES γδ TCR, if the TCR binds analogously to αβ TCR/CD1 complex enabling lipid discrimination. A number of additional residues were mutated on the opposite face of EPCR.*

In these constructs, solvent-exposed amino acids were mutated to alanine residues to determine whether these substitutions affected LES γδ TCR binding to EPCR. Many of the constructs generated contained a single amino acid mutated to alanine, however several constructs were generated in which multiple residues in close proximity were mutated in combination, in order to more rapidly identify important residues. If any of these compound mutants highlighted a region of interest, then individual mutant EPCR constructs could be generated to focus on the residues of interest (Figures 4.8/4.9). In total, twelve amino acids on the α1 helix were mutated in eight constructs and eleven amino acids on the α2 helix were mutated in nine constructs (Table 4.2). In addition, as the α1/α2 platform of EPCR was predicted to be tethered to plasma membrane by a flexible stalk, a number of amino acids on the β-sheet or the inter connecting strand loops were also mutated to alanine, as this region could be accessible for interaction with the LES γδ TCR.

4.3 Site directed mutagenesis of EPCR, transfection and large scale expression

Wild type EPCR (wtEPCR) in the pMTBiPBirAHis vector was used as the DNA template in the Site Directed Mutagenesis (SDM) PCRs. After successful PCR and verification of the point mutations by DNA sequencing (Figure 4.10) the mutated constructs were introduced into S2 cells by calcium phosphate transfection. S2 transfectants were selected for hygromycin resistance as previously described in Chapter 3. Expression of EPCR mutants following induction by copper sulphate was confirmed by SDS-PAGE and western blotting (Figure 4.11). After verification of mutant EPCR expression by western blot analysis, large scale production was undertaken to produce and purify the mutant EPCR proteins, as described for the soluble TCRs in Chapter 3. Ni-NTA purified EPCR mutants were biotinylated utilising BirA enzyme and free biotin was removed by gel filtration. SDS-PAGE analysis of biotinylated protein indicated that the level of purity was sufficient for SPR analysis (Figure 4.12).

4.4 Effect of α1/α2 mutations on EPCR recognition

SPR analysis of EPCR mutants was performed as described in Figure 4.13a. In each experiment wtEPCR was immobilised alongside mutant EPCR and an irrelevant control protein, HLA-A2. Equilibrium binding responses were measured over a range of LES γδ TCR concentrations and the apparent Kd was determined for the wild type and mutant in each individual experiment. Representative binding plots are shown for EPCR S67A and Q149A compared to wtEPCR in a single experiment (Figure 4.13b). These residues are located on the α1 and α2 helix, respectively, and bound to LES γδ TCR with a similar affinity as wtEPCR (Figure 4.13c). In addition, both S67A and Q149A mutants bound the 2E9 mAb equivalently to wtEPCR (Figure 4.13d). Analysis of other mutants on the α1/α2 helices w

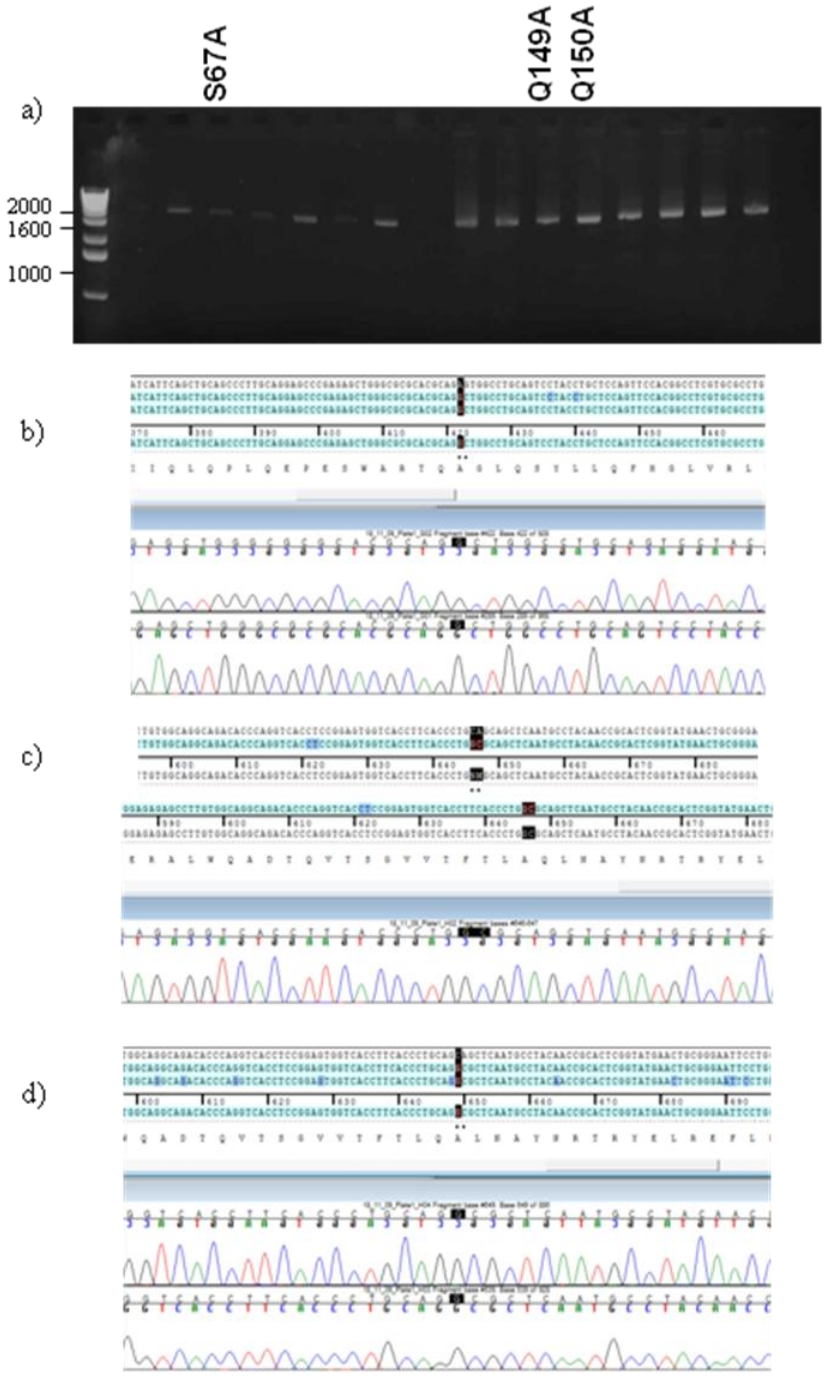


Figure 4.10 Site-directed mutagenesis of EPCR. a) Gel electrophoresis analysis of the products of site-directed mutagenesis PCRs of the α1 and α2 helices of EPCR loaded on a 1% agarose gel. Lane 1 loaded with 1Kb DNA ladder, lanes 2 to 18 loaded with all 17 SDMs to mutate solvent-exposed residues on both helices. b-d) Sequencher analysis of wtEPCR sequence compared to EPCR mutants S67A (b), Q149A (c) and Q150A (d). Nucleotides highlighted by black dots indicate the bases that have been changed to introduce the Alanine mutation.

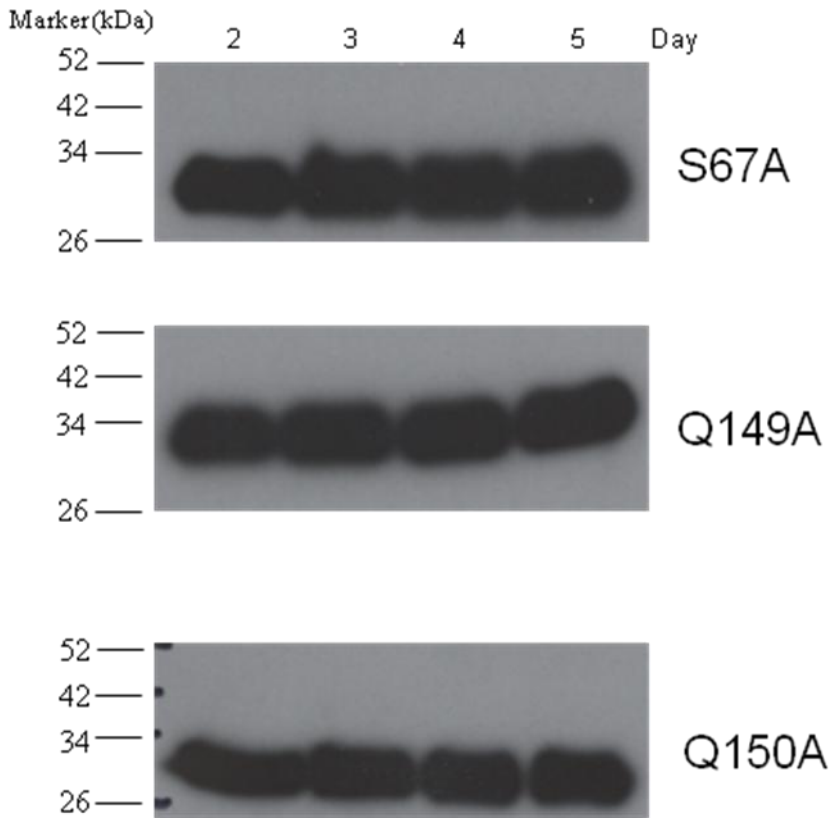


Figure 4.11 Western blot analysis to test for expression of mutant EPCR constructs. Mutated EPCR constructs were transfected into S2 cells and selected for hygromycin resistance. Cells were induced to express protein by addition of copper sulphate (0.5 mM), and cell supernatants were harvested on the indicated days and analysed by SDS-PAGE and western blotting. Expression of His-tagged EPCR mutants was detected by blotting with anti-His tag mAb, goat anti-mouse IgG-HRP, and ECL.

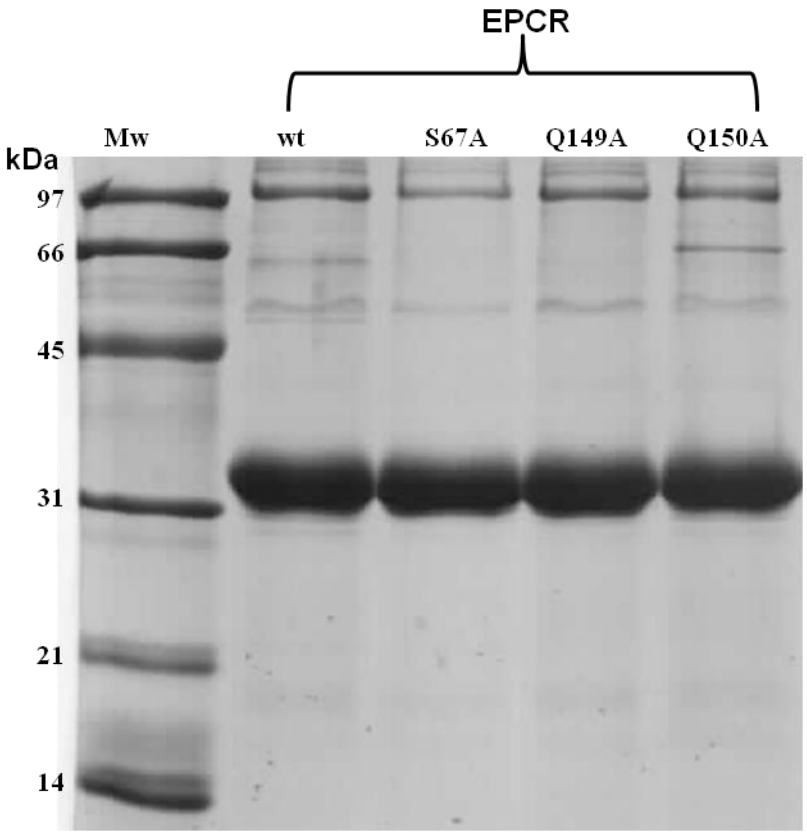


Figure 4.12 SDS-PAGE analysis of purified biotinylated wild type and mutant EPCR proteins utilised in SPR experiments. wtEPCR and EPCR mutants were purified from S2 cell supernatant by nickel chelate chromatography, biotinylated and further purified by size exclusion. Purified proteins were analysed by SDS-PAGE and Coomassie staining prior to SPR studies.

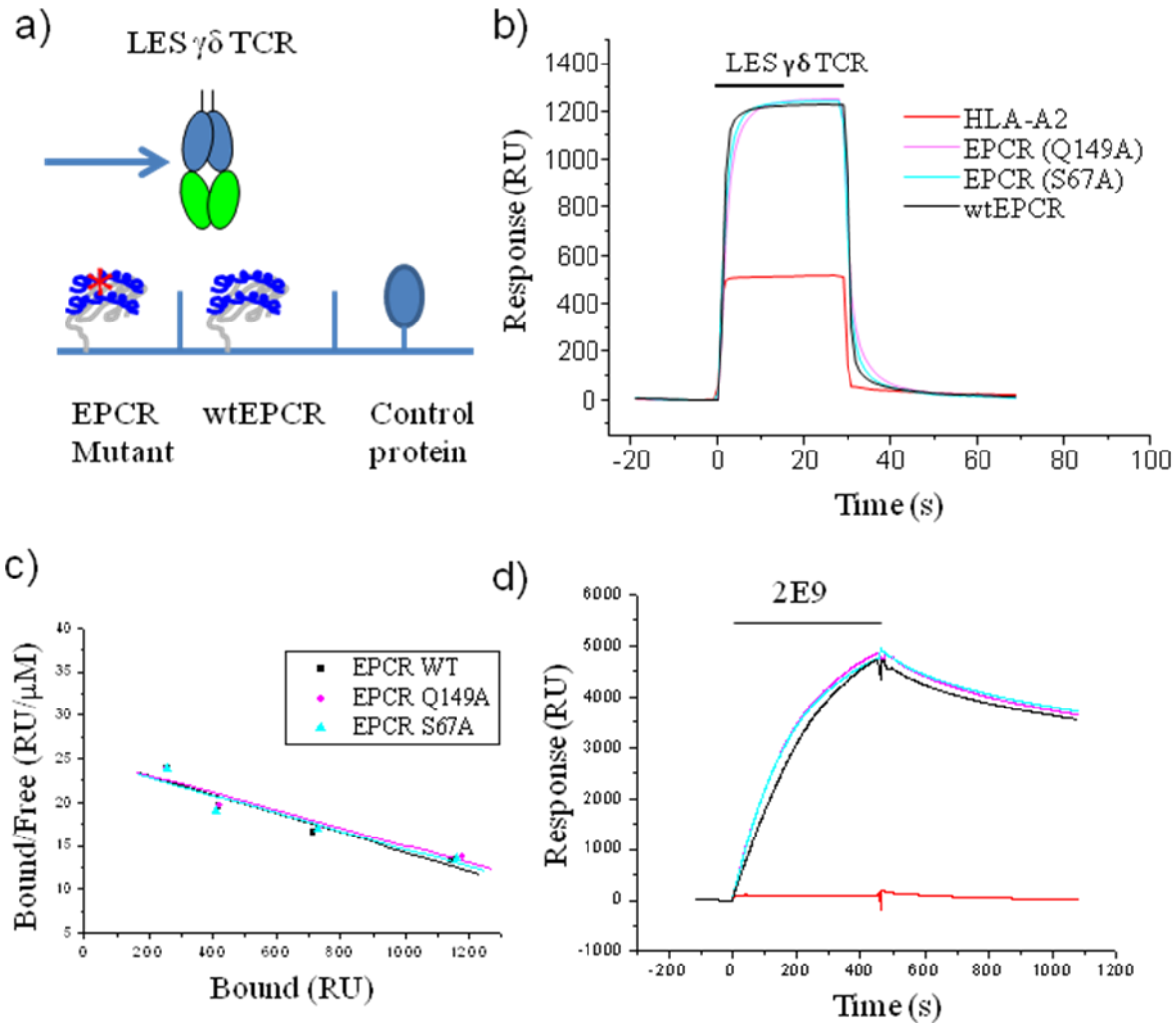


Figure 4.13 Analysis of EPCR mutants using SPR. a) Cartoon of experimental set-up, showing biotinylated wild type and mutant EPCR, as well as control HLA-A2, were immobilised on a streptavidin-coated CM5 chip and binding of soluble LES γδ TCR was measured. b) SPR profile of LES γδ TCR (43 mM) injected over immobilised control (HLA-A2, red, 4291 RU), wtEPCR (black, 4085 RU), and EPCR mutants S67A (blue, 4089 RU) and Q149A (pink, 4047 RU). c) Scatchard plot of the binding data for wtEPCR (Kd=90.9 mM) and mutants S67A (Kd=97.1 mM) and Q149A (Kd=99.1 mM). d) 2E9 mAb (2 mg/ml) was then injected over the same proteins, and response measured.

performed, and the binding affinity of wtEPCR relative to mutant EPCR was plotted (Figure 4.14; Table 4.3). All α1/α2 mutants bound LES γδ TCR with similar affinities, with the exception of mutations at E86 and R87 which had a marginally higher affinity for the TCR. This is particularly interesting as E86A/R87A is present on the α1 helix and is sited at the membrane distal region of EPCR, on the assumption that the long axis of the EPCR ectodomain (parallel to the α1 and α2 helices) is perpendicular to the plane of the membrane. This is consistent with a model whereby LES γδ TCR would dock to the membrane distal region on EPCR. However, mutants T145A, F146A and Q149A are also prominent at the membrane distal region on the α2 helix and highlighted no change in LES γδ TCR recognition. The E86A/R87A mutations substituted large charged side chains for small alanine side chains and could have affected the overall structure of EPCR, supported by evidence that 2E9 mAb binding to this mutant is decreased. Several other α helical mutations revealed reduced binding of 2E9 (Table 4.3), whereas the R81A/L82A mutant combination completely abolished 2E9 binding, suggesting these residues are critical for 2E9 recognition.

α1 helix	**equilibrium Kd mutant (μM)**	**equilibrium Kd wt (μM)**	**wt Kd/mutant Kd**	**2E9 binding**	**Error +/-**
E58 & E60	160.2	184.5	1.15	+	19.57
R64	33	29.5	0.89	+	5.29
S67	100.5	97.8	0.97	+	32.07
S71 & L74	109.9	104.6	0.95	+	21.78
Q75	280.2	279.9	1.00	+	75.91
R81 & L82	184.4	177.9	0.96	-	12.76
Q85	157.9	136.1	0.86	+/-	32.5
E86 & R87	77.2	138.1	1.79	+/-	14.34
α2 helix					
T145	139.4	177.9	1.28	+	10.62
F146	28.7	29.5	1.03	+/-	3.77
Q149	105.7	97.8	0.93	+	25.12
Q150	91.2	83.2	0.91	+	15.76
Y154	133.5	122.9	0.92	+	26.0
R156	32.8	28.7	0.88	+	5.86
E160	135.1	134.4	0.99	+	10.52
E163	91.7	117.7	1.28	+	6.68
D167,Q171 &	83.9	105.7	1.26	+/-	30.4

K175					

Table 4.3 LES γδ TCR binding affinity to EPCR α1/α2 mutants. *Point mutations on the α1 and α2 helices of EPCR were analysed for LES γδ TCR binding by SPR. Equilibrium binding analyses were performed for mutant EPCR and wild type EPCR in a single experiment and apparent dissociation constants (Kd) were calculated. The ratio of wild type Kd/mutant Kd was calculated in each experiment. A ratio of ~1 indicates equivalent binding. Mutations in the α-helices of EPCR indicated this region was not involved in LES γδ TCR binding. EPCR mutants were also analysed for 2E9 mAb binding to confirm that these mutations did not affect overall conformation of EPCR. Where Kd values were recorded, errors in the Kd values (ie Kd +/- x μM) were determined from estimates of the standard error from the Equilibrium binding as calculated in Origin V3. In each case, the range given (ie +/- x μM) represents the largest deviation from the estimated Kd value. Ultimately, determination of the binding mode of the LES TCR/EPCR interaction relied entirely on mutations that abolished binding, for which Kd estimation was clearly not possible.*

Strikingly, these results indicated that it is unlikely LES γδ TCR binds EPCR in a similar manner to αβ TCR/CD1 interaction, as lipid discrimination would require interaction with the α1/α2 helical residues. However, analysis of the β-sheet mutations highlighted several residues which affected TCR binding (Table 4.4). For instance, EPCR mutants N119A and S121A failed to bind TCR but interestingly, these residues make up one of the glycosylation sites. It is unclear whether the N119A and S121A mutants are locally misfolded due to the disruption to the glycosylation motif during protein synthesis and folding in the endoplasmic reticulum, as mutant S121A also binds 2E9 mAb with decreased affinity. Furthermore,

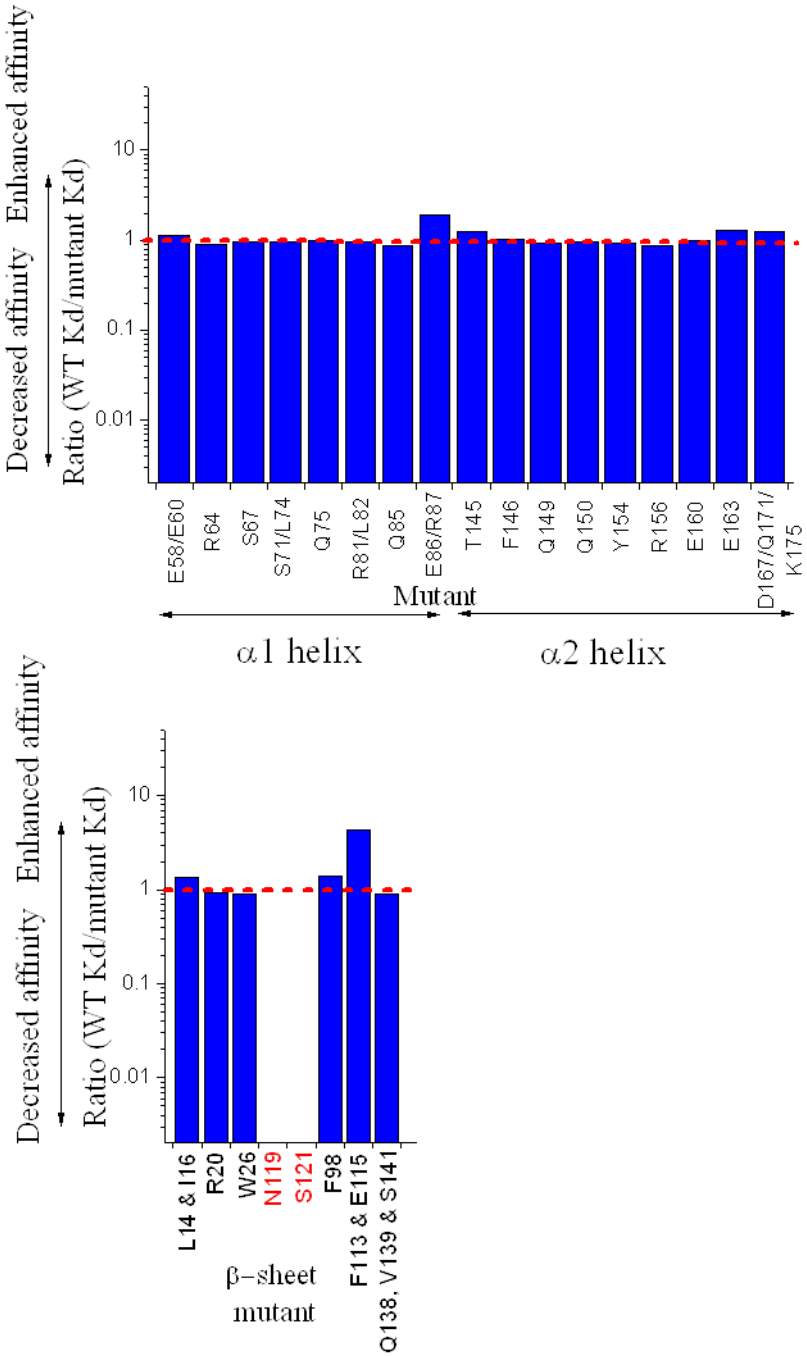


Figure 4.14 LES γδ TCR binding affinity for EPCR mutants (mutant Kd) with alanine substitution of α1 or α2 helix residues relative to EPCR wild type affinity (WT Kd) measured in the same experiment (a). Binding affinity for EPCR mutants with alanine substitution on β-sheet residues relative to EPCR WT (b). Red dashed line indicates an equal binding affinity.

double mutant F113A/E115A displayed a four-fold increased TCR binding. The increase in TCR binding may have resulted from the longer side chains perturbing the LES γδ TCR-EPCR interaction. This was particularly interesting as these amino acids were positioned on the β-sheet and away from the lipid binding pocket of EPCR, highlighting a novel region on the β-sheet (Figure 4.15) as a potential TCR binding site. To further examine this hypothesis, additional mutations on the β-sheet surrounding residues 113-121 were generated and analysed.

β-sheet	equilibrium Kd mutant (μM)	equilibrium Kd wt (μM)	wt Kd/mut Kd	2E9 binding	Error +/-
L14A & I16A	137.3	184.5	1.34	+	66.16
R20A	34.5	32.4	0.94	ND	1.75
W26A	32	28.7	0.90	+	5.02
N119A	0	131.6	0	+	0
S121A	0	56.3	0	+/-	0
F98A	88.2	122.9	1.39	+	28.05
F113A & E115A	27.7	117.7	4.25	+	11.7
Q138A,V139A & S141A	135.1	123.4	0.91	+	12.63

Table 4.4. β-sheet mutations scanning the platform region of EPCR. *Effect of EPCR β-sheet mutations on LES γδ TCR binding. Equilibrium binding analyses were performed and Kd calculated for WT and mutant EPCR in each experiment. The ratio of WT Kd/mutant Kd was calculated. A ratio of approximately 1 indicates equivalent binding of LES γδ TCR to WT and mutant EPCR. Ratio of >1 indicates increased binding, while 0 indicates no binding was observed to mutant EPCR. 2E9 was also used to confirm EPCR structure was not disrupted by mutagenesis. Note that 2E9 binding to EPCR-R20A was not determined (ND).*

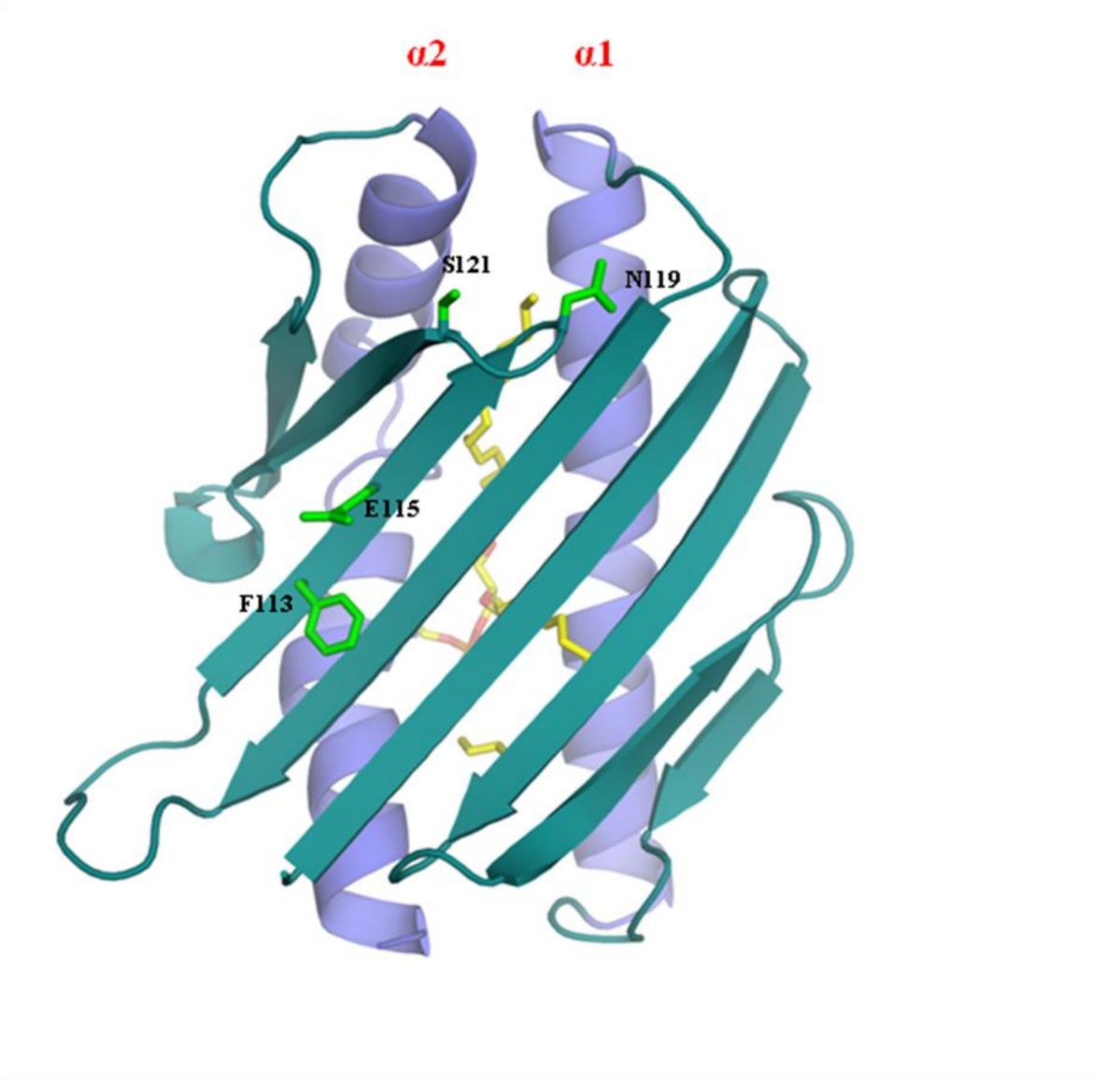


Figure 4.15 Amino acid residues on EPCR indicating areas potentially involved in LES γδ TCR binding. Highlighted in green sticks are residues F113/E115, mutation of which results in a higher affinity for TCR binding, and N119 and S121, mutation of which abrogates LES γδ TCR binding.

4.5 Identification of key residues involved in LES γδ TCR binding to EPCR

On examining the EPCR structure to identify solvent exposed amino acids on the β-sheet in close proximity to residues 113–121, an additional 27 amino acids were selected for SDM analysis (Table 4.5). Firstly, to determine the precise role of F113 and E115 with respect to LES TCR binding, each residue was separately mutated to alanine. In addition, the nearby residue A117 was mutated to a smaller amino acid, glycine (A117G), as well as a large charged residue, aspartic acid (A117D). This would highlight whether the beta sheet residues represented a *bona fide* TCR binding site. Moreover, to determine if S121 was truly involved in TCR binding, it was mutated to threonine to preserve the glycosylation site. Finally, to assess the limits of the proposed TCR binding site, several amino acids on neighboring β-strands and loops were mutated to alanine, as single residue mutants or in pairs (Figure 4.16).

Each mutant was analysed for LES γδ TCR binding by SPR as described earlier, and compared directly to wtEPCR in each experiment. A ratio of wild type Kd/mutant Kd was calculated and a number of residues emerged which affected TCR binding (Figure 4.17). The most striking observation consisted of individual mutants F113A and E115A abolishing LES γδ TCR binding, confirming that these residues were of interest. This result was contrary to the combined double mutations on EPCR F113A/E115A which displayed a 4-fold increase in LES γδ TCR binding. However, one caveat is that for F113A, the lack of LES TCR recognition was accompanied by a decrease in 2E9 binding, and the loss in TCR binding by this mutation resulted from local misfolding or altered conformation of EPCR. However, E115A exhibited a loss in LES γδ TCR binding accompanied by normal levels of 2E9 binding indicating correct conformation in EPCR, confirming likely involvement in the interaction. S121T bound LES γδ TCR but with a slightly decreased affinity, and nearby

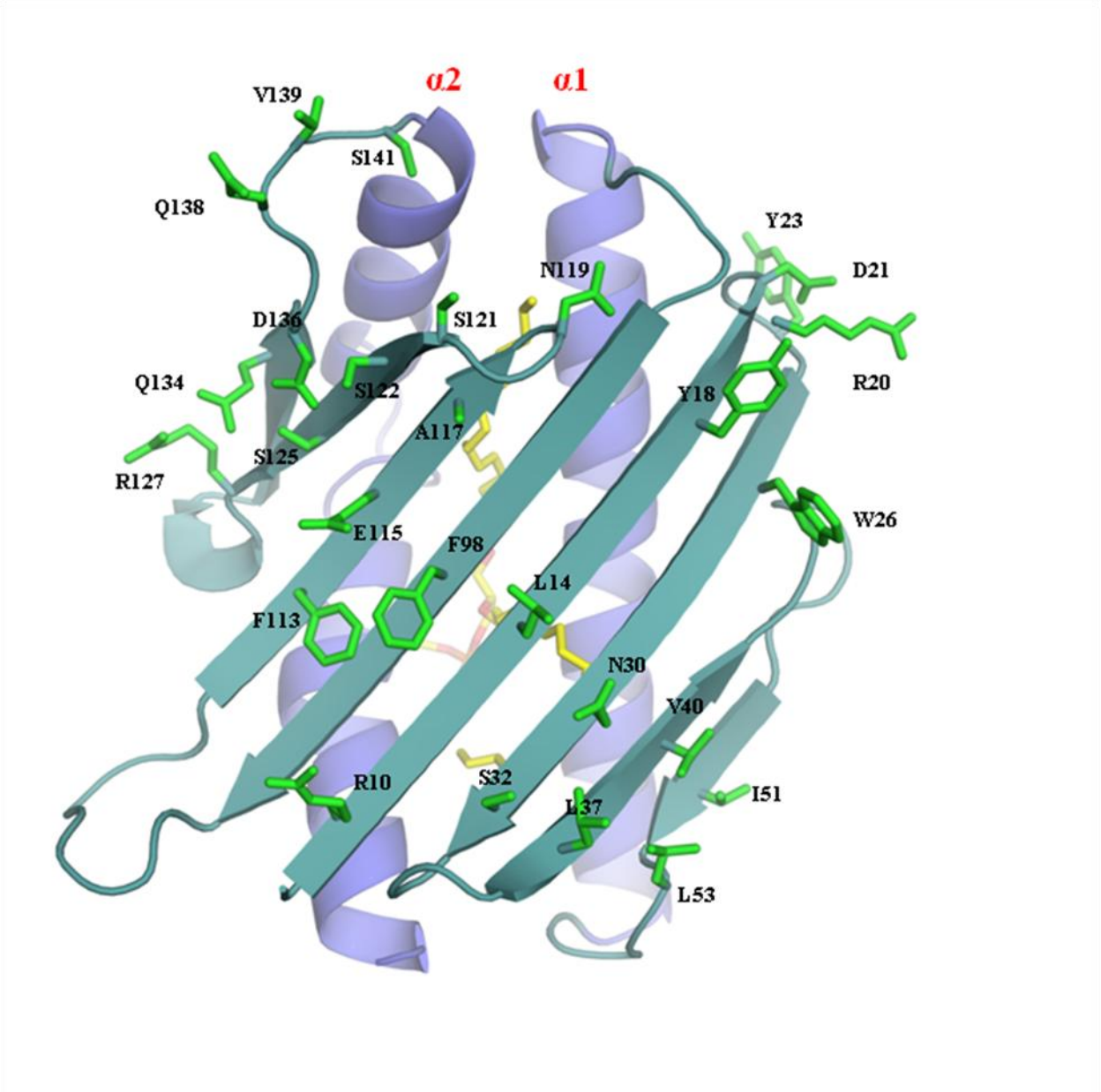


Figure 4.16 Further amino acids mutated on the β sheet to identify the mode of LES γδ TCR binding to EPCR. Highlighted in green sticks are the side chains of solvent exposed residues which were most prominent and are therefore potentially involved in TCR binding.

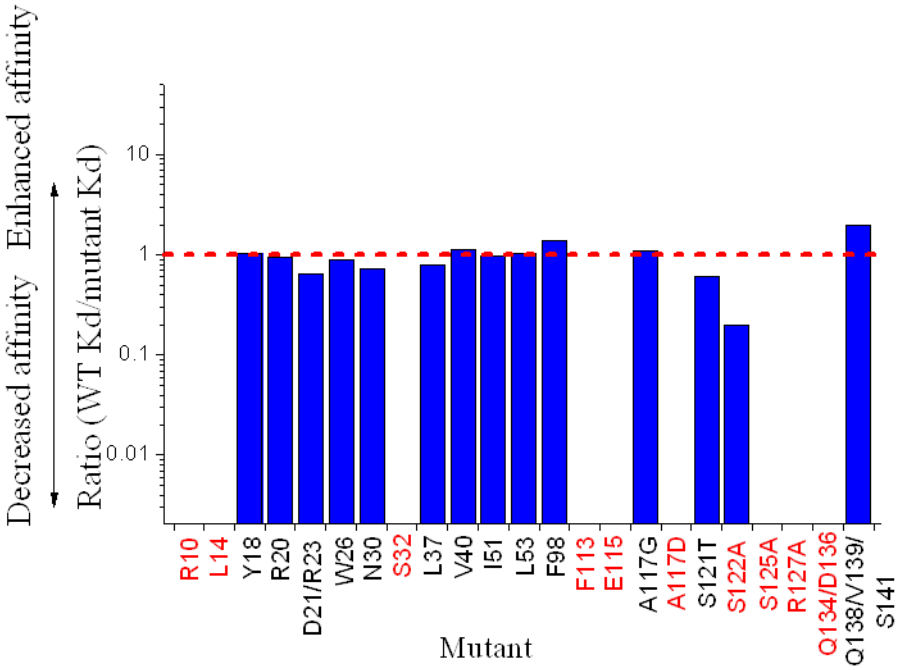


Figure 4.17 Analysis of further EPCR β-sheet mutations by SPR, showing the ratio of WT LES γδ TCR binding affinity to mutant of solvent exposed mutant constructs. Data plotted by normalising affinity of wtEPCR tested against each mutant. Indicated in red are the residues mutated which demonstrated a decrease in LES γδ TCR affinity for EPCR.

S122A reduced the affinity by 5-fold. The fact that S121T and S122A demonstrate decreased affinities despite the preservation of the glycosylation motif, suggests the serine side chains themselves, rather than the glycosylation site itself, were important in LES γδ TCR binding. The A117 residue seemed also to be involved in binding, as mutagenesis to a larger aspartic acid inhibited binding, while mutation to the smaller glycine revealed no effect. A number of residues on neighboring β-strands also affected TCR binding, and the majority of these bound 2E9 mAb equivalently to wtEPCR, suggesting the mutation did not alter the overall structure of EPCR. In addition, residue S32 is part of an N linked glycosylation site at N30 and substitution to alanine disrupted TCR binding. However, this disruption did not seem to be due to absence of the glycosylation site or misfolding in the absence of the glycosylation, as the N30A mutant bound LES γδ TCR equivalently to wild type and both bound 2E9 mAb. In all, 12 residues on the β-sheet affected binding (Table 4.5), and an extensive binding site emerged on the β-sheet face opposite to the lipid binding platform (Figure 4.18).

2nd round mutants	**KD equilibrium mutant (μM)**	**KD equilibrium wtEPCR (μM)**	**WT Kd/mutant Kd**	**2E9 binding**	**Error +/-**
R10A	0	51.8	0	+	0
L14A	0	84.8	0	+	0
Y18A	88.3	91.4	1.04	+	16.17
D21A/Y23A	59.5	38.5	0.65	+/-	10.76
N30	102.5	74.5	0.73	+	12.7
S32A	0	74.5	0	+	10.8
L37A	64.9	51.8	0.80	+	1.77
V40A	74.3	83.2	1.12	+	10.63
E42A	290	138.1	0.48	+	166.14
N47A	34.6	86.1	2.49	+	6.55
T49A	75.3	94.6	1.26	+	10.8
I51A	77.7	75.6	0.97	+	5.29
L53A	175.3	182.5	1.04	+	28.4
T94A/R96A	79.6	49	0.62	+	24.88
F113A	0	105.7	0	+/-	0
E115A	0	140	0	+	0
A117G	77.4	86.1	1.11	+	22.11
A117D	0	67	0	+	0
S121T	37	22.6	0.61	+	13.74

S122A	717	140	0.20	+	166
S125A/R127A	0	91.4	0	+	0
S125A	0	84.8	0	+	0
R127A	0	182.5	0	+	0
Q134A/D136A	0	49	0	+	0
N155A	78.1	80.1	1.03	+	15.13
T157A	59.4	38.5	0.65	+	14.19

Table 4.5. LES γδ TCR binding affinity data of further selected residues for mutagenesis. *Highlighted are the affinity of each β-sheet mutation compared to the wtEPCR, 2E9 antibody analysis indicates conformation of EPCR.*

4.6 2E9 binding site identification

Binding analyses of the 2E9 antibody to each mutant highlighted several other amino acids which affected 2E9 binding. Combined with the previously identified residues on the α1/α2 helices, it is predicted that the main 2E9 binding site encompasses residues R81, L82, Q85, E86, E87 and F146 (Figure 4.19). Further, other residues that may contact 2E9 include D21 and Y23 which are located on the side of platform. It is unclear how F113A and S121A on the underside of the platform can affect 2E9 binding; potentially these mutations could affect folding of the molecule, causing a localized change in conformation of EPCR that may have an effect on 2E9 binding to other regions of the molecule. Likewise, mutation of D167A/Q171A/K175A also demonstrated reduced 2E9 binding. These residues are on the C-terminal end of the α2 helix, located distal to the putative 2E9 binding site. It is possible that mutation of these residues may cause a conformational change in the α2 helix such that it affects 2E9 binding (Figure 4.19).

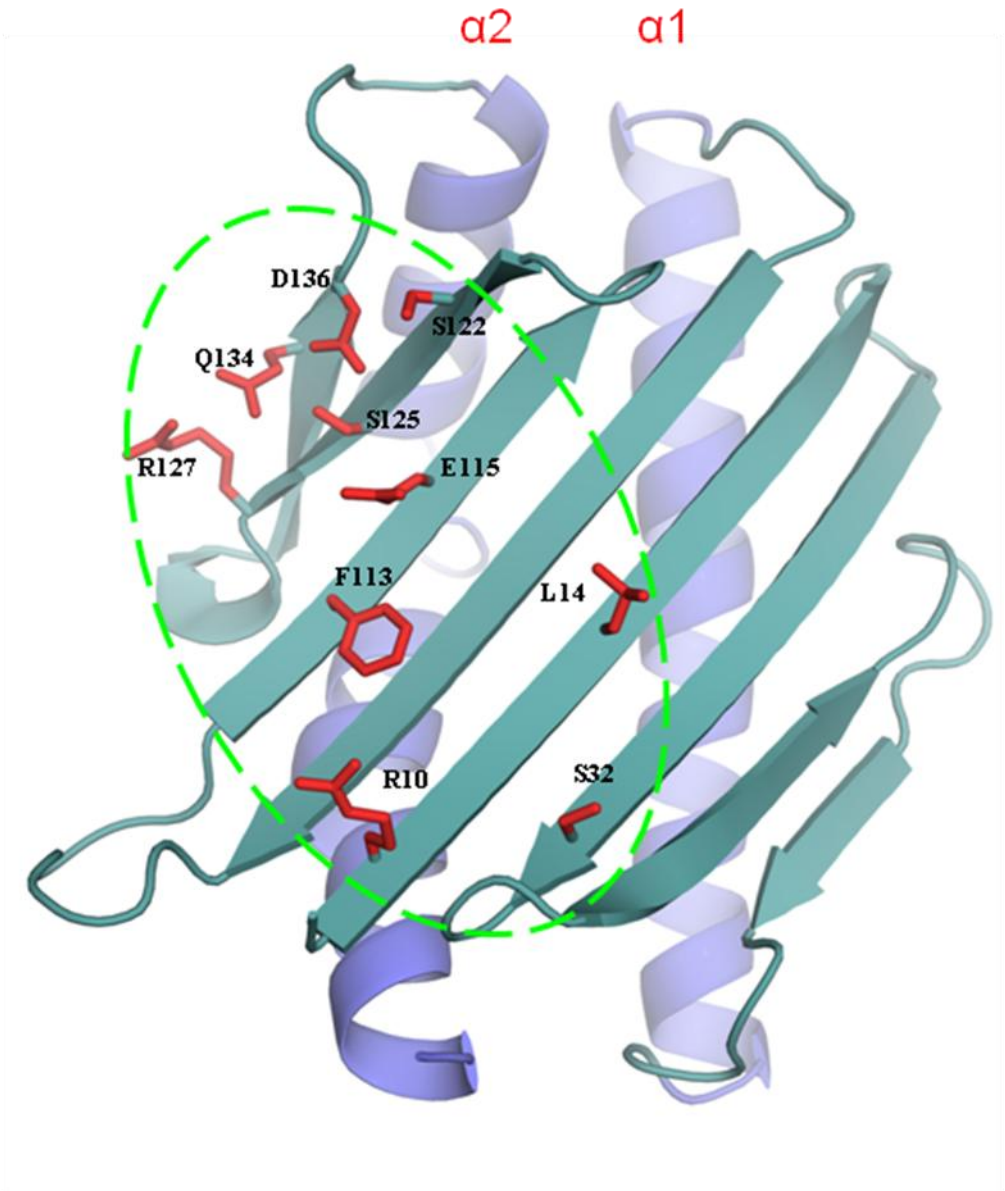


Figure 4.18 LES γδ TCR binding mode to EPCR. Depicted in red are the residues which, when mutated to alanine, demonstrated decreased LES γδ TCR binding to EPCR. This diagram highlights a region on the β-strand platform where the LES γδ TCR may dock to the EPCR protein.

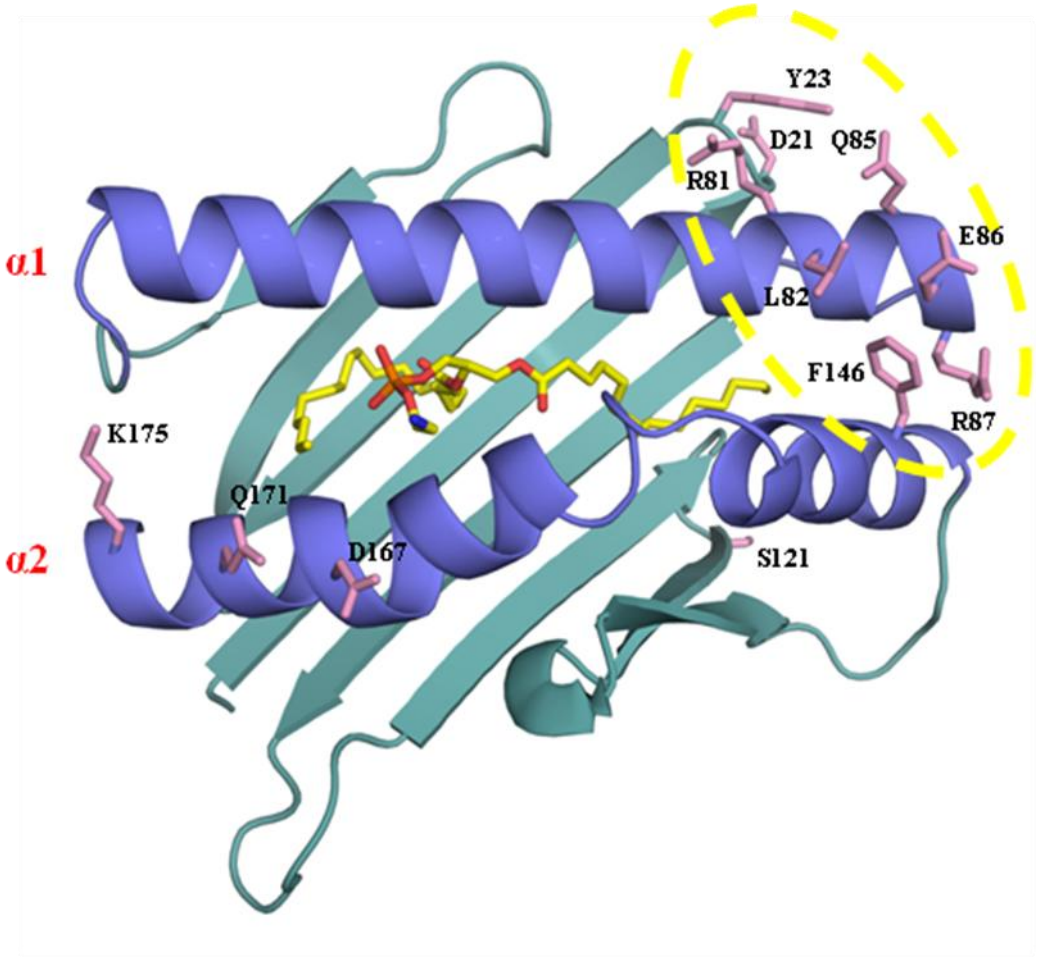


Figure 4.19 2E9 mAb binding mode to EPCR. Depicted in pink are the residues, which when mutated to alanine highlighted a change in 2E9 binding to EPCR. Shown in yellow dotted line is the 2E9 binding site. Indicating 2E9 antibody interacts with the α1/α2 helices of EPCR, away from the proposed LES γδ TCR docking site on the β-sheet.

In summary, the majority of residues which affected 2E9 antibody binding were located on one end of the α1/α2 platform of EPCR. The 2E9 binding site was distinct from the LES γδ TCR recognition site, suggesting 2E9 did not directly prevent TCR binding in functional assays (Figure 4.18). Rather, because 2E9 is a pentameric IgM, binding of 2E9 sterically inhibited binding of LES γδ TCR to EPCR. Interestingly, a number of residues which affect 2E9 binding were identified as critical for APC binding (Oganesyan *et al,* 2002) suggesting 2E9 could be a good blocking antibody for the EPCR-APC interaction (Figure 4.20).

4.7 Crystallisation trials of the LES γδ TCR-EPCR complex

To enable a greater molecular understanding of the LES γδ TCR-EPCR interaction crystallization of the complex was attempted. Expression and purification of LES γδ TCR and EPCR was performed as described in Chapters 3 and 4. After concentration of the LES γδ TCR and EPCR, the purity of the proteins was evaluated by SDS-PAGE analysis (data not shown). This highlighted that the proteins were of acceptable purity for subsequent crystallisation complex trials. SPR analysis of LES γδ TCR-EPCR interaction revealed that the interaction was within the micro-molar range similar to αβ TCR-pMHC affinities (1-100 μM). Despite the low affinity observed for the αβ TCR-pMHC interactions, several of these ternary complexes have readily crystallized allowing their structures to be determined (Rudolph *et al,* 2006), thereby suggesting that LES γδ TCR-EPCR complex is a valid candidate for structural studies. Initial complex crystallization trials were attempted using the hanging drop vapour diffusion method at a concentration of 5.3 mg/ml. These trials were unsuccessful presumably due to the fact that the majority of the hanging drops were under-saturated.

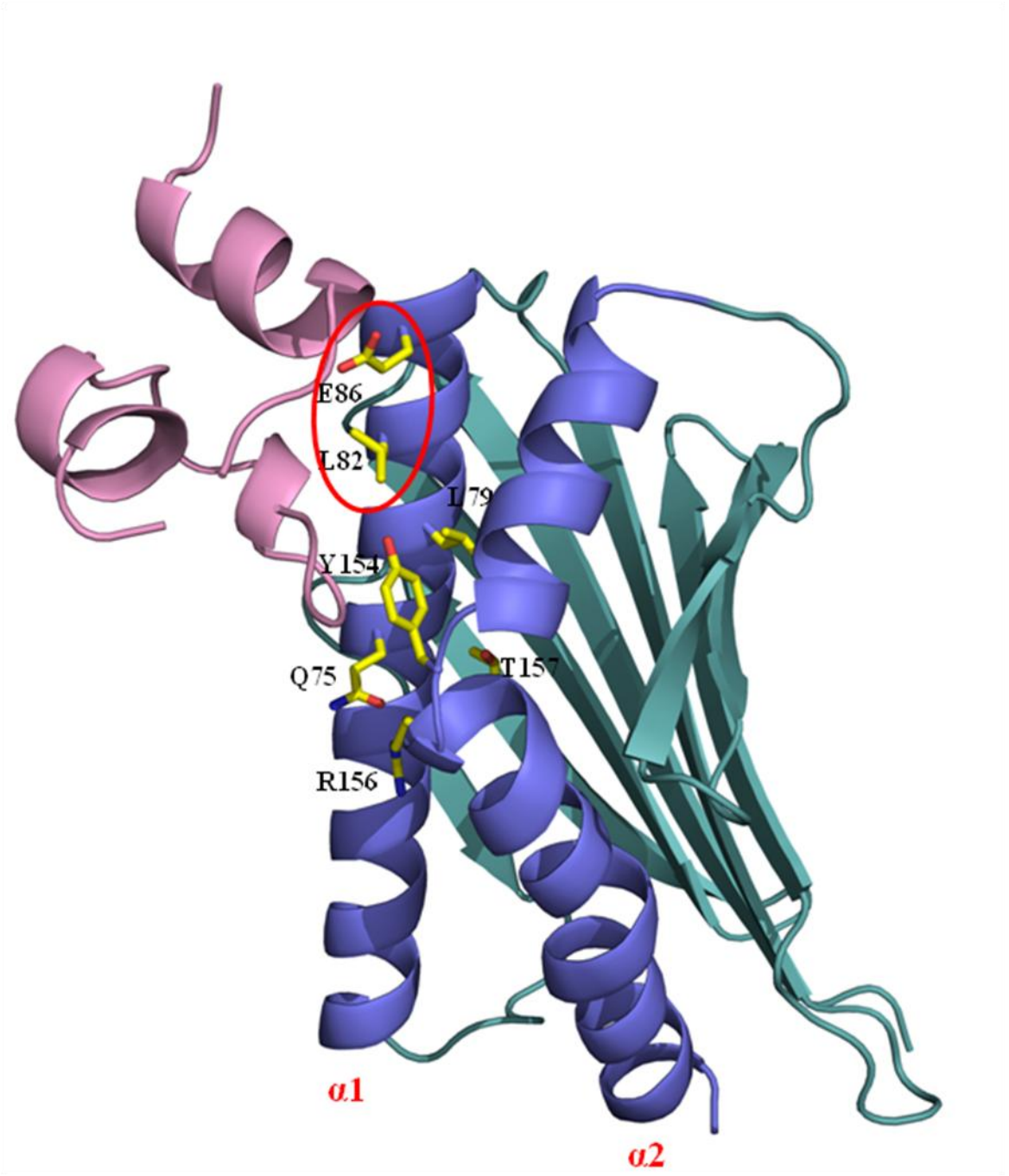


Figure 4.20 Crystal structure of EPCR in complex with the Gla domain of Protein C (Oganesyan *et al*, 2002). The Gla domain of protein C (depicted in pink) interacts with the α1/α2 helices of EPCR (blue), away from the proposed LES γδ TCR docking site on the β-sheet. Residues of EPCR that interact with Gla domain are highlighted in ball and stick format and labelled. EPCR residues that are circled in red are involved in 2E9 binding, indicating 2E9 mAb binding may inhibit protein C docking to EPCR.

Further attempts for setting up crystallization trials at a higher concentration were limited due to problems associated with producing highly concentrated amounts of LES γδ TCR. Also, previous structural studies of EPCR demonstrated the region following the α2 helix that links EPCR to the transmembrane domain is very flexible. It is possible that this region may have prevented efficient crystallisation. To overcome this issue and potentially facilitate crystallization of the complex, a new construct was designed by Dr Carrie Willcox which lacked the flexible region at the C-terminus of the protein, but still encompassed a His tag for purification. Unfortunately, the new complex trials with the shortened EPCR construct also revealed a lack of viable crystallization hits. This was attributed to the drops being under-saturated, despite the fact that the trials were set up at a protein concentration of 10 mg/ml.

The final strategy involved utilizing a high affinity mutant EPCR construct (F113A/E115A; Kd= ~25 μM) previously identified by SPR analysis. Notably, the mutant construct also lacked the flexible stalk region. It was possible that the increased affinity of the LES γδ TCR-EPCR complex may facilitate crystal formation. However, the mutant EPCR high affinity construct proved to be unstable when concentrating to high levels. To increase the solubility of the high affinity EPCR mutant construct, a thermal stability assay was performed, and a buffer condition was identified that appeared to increase stability. However, when EPCR was then concentrated in this optimized buffer, the stability of the protein, as judged by its propensity to aggregate, was observed to be similar in the optimized buffer at high concentrations, and only a limited number of crystallization complex trials were set up.

4.8 Discussion

Although several molecules have been proposed to stimulate human γδ T cells, none have been demonstrated to bind the γδ TCR in direct binding studies. EPCR has been identified as a ligand for human Vγ4Vδ5 TCR derived from a T cell clone which recognized tumour cell lines and CMV infected cells. Although the affinity of this interaction is low, it is within the range of affinities for αβ TCR-pMHC studied to date (1-100μM) (Rudolph *et al*, 2006). Notably, it is much lower than the only γδ TCR interaction known, that of the murine G8 TCR for non-classical MHC molecules T10/T22 (0.1 μM) (Adams *et al*, 2005). It is plausible that similar to αβ TCRs, γδ TCRs may display a broad range of affinities for their ligands.

Further SPR analysis revealed LES γδ TCR binds a conformational epitope on EPCR. Previous analysis of TCR binding had demonstrated EPCR was necessary but not sufficient for activation of the LES γδ T cell clone or the transfected JRT3-LES cell line. This could have been due to intrinsic or extrinsic factors associated with EPCR. EPCR-intrinsic factors could include effects of glycosylation of the protein, as EPCR contains four N-linked glycosylation sites within the primary protein sequence. Indeed, analysis of tumour cell lysates indicated EPCR was heavily and differentially glycosylated in different cell lines (Willcox *et al*, 2012). However, there was no obvious correlation between the extent of EPCR glycosylation and ability of the cell line to activate LES γδ T cells. To directly address this, recombinant EPCR was deglycosylated by treatment with PNGase F. Purified EPCR was completely deglycosylated as determined by SDS-PAGE, but still bound LES γδ TCR with the same affinity as glycosylated EPCR. Therefore, glycosylation of EPCR is unlikely to explain the observed differential reactivity.

Another EPCR-intrinsic factor could include differential membrane localization of EPCR in activating and non-activating cells. The short cytoplasmic tail of EPCR contains a cysteine

which can be palmitoylated (Xu *et al*, 1999). Palmitoylation is a reversible post translational modification in which a palmitic acid chain is attached to juxtamembrane cysteine residues and affects subcellular trafficking and protein association (Linder and Deschenes, 2007). In endothelial cells, palmitoylation of EPCR is thought to result in localization of EPCR to caveolae. Cell based studies carried out in our laboratory have shown EPCR staining in tumour cell lines to be highly punctuate consistent with localization of EPCR to lipid rafts or caveolae (Appendix I, Figure 1d). However, this is observed in both activatory and non-activatory cell lines, and hence is unlikely to explain the pattern of reactivity to $2E9^{+}$ cells.

A third possibility is that EPCR has the capacity to present a stress lipid to the TCR. This is supported by the fact that EPCR is structurally related to CD1d, a lipid presenting molecule recognized by NKT cells. Strikingly, EPCR has also been demonstrated to possess a lipid molecule bound between the helices of the protein. The hypothesis that EPCR presents a lipid to the γδ TCR was examined in several ways. Cell lines which activate or fail to activate the LES γδ TCR were transduced to express sEPCR, which was purified and analysed for LES γδ TCR binding by SPR. No differences in affinity for TCR were observed between sEPCR produced in activating (U373) or non-activating (HeLa) cells, or between uninfected and CMV infected U251 cells (Appendix I Supplementary Figure 4). However, EPCR molecules that lack the transmembrane and cytoplasmic tail may not traffic throughout the cell in the same manner as native EPCR and hence are unlikely to enter the correct organelle to be loaded with appropriate lipid. Therefore, alanine scanning mutagenesis was used, with the intention that mutations on the α1/α2 helices would affect TCR binding if bound lipid was being discriminated. However, the α1/α2 helical mutations showed no change in LES γδ TCR binding suggesting that the TCR does not to discriminate lipids presented by EPCR. Following additional mutational analysis the region of LES γδ TCR binding was identified on the β-sheet face beneath the lipid binding platform of EPCR. This region would be accessible

for TCR binding as the EPCR structure predicted the α1/α2 platform as being attached to a flexible stalk, rather than tethered in a horizontal position like class I MHC or CD1d by interactions with α3 and β_2M. This flexibility may allow the LES γδ TCR direct access to the binding site on the β-sheet platform. Therefore, the difference in recognition of activatory and non-activatory cell lines does not seem to be related to an EPCR-intrinsic factor. Instead, LES clone and Jurkat transfectant recognition of activating tumour lines and CMV infected cells has been shown to be highly dependent on costimulatory signals including CD58-CD2 and CD54-CD11a. Despite these observations, blocking these interactions led to incomplete inhibition of activation (Appendix I Figure 7), suggesting other yet to be identified costimulatory complexes may be involved.

Several alloreactive murine γδ clones have been identified which recognize classical MHC or MHC-like molecules (Crowley *et al*, 2000; Hampl *et al*, 1999, Schild *et al*, 1994), but the two clones which have been studied extensively have been shown to interact with residues on the α1/α2 helices, similar to αβ TCR-pMHC interactions. The unique binding mode of LES γδ TCR for EPCR may be indicative of γδ T cells ability to recognize intact self stress antigens in an antibody-like manner. This is reminiscent of CD4/CD8/MHC-I$^-$/II$^-$ deficient mice, in which αβ T cells are not MHC-restricted, but recognized intact self antigen (Tikhonova *et al*, 2012).

To gain greater molecular insights into the LES γδ TCR-EPCR interaction, co-crystallisation of the two proteins was attempted, but has thus far been unsuccessful. To overcome intransigent LES γδ TCR-EPCR complex crystallization, a number of strategies used in previous structural studies could be attempted. These involve decreasing the level of protein glycosylation by expressing proteins in the presence of tunicamycin (Adams *et al*, 2005), or treatment of purified sEPCR and LES γδ TCR with PNGase F or endoglycosidase H to remove N-linked glycosylations. One could also try to increase the concentration of LES γδ

TCR and EPCR in complex crystal trials to enhance protein nucleation. Furthermore, introduction of an additional domain to the complex may encourage crystallization, by altering crystal contacts. This could include addition of an Ig-domain to the C-terminus of EPCR, or addition of APC to the complex trials.

Structural analysis of the LES γδ TCR-EPCR complex will undoubtedly complement the mutational studies and help validate the proposed binding site identified in this chapter. In addition, the complex structure may elucidate other residues on EPCR which interact with LES γδ TCR, as well as determine which regions of the TCR are vital for EPCR binding. It will be interesting to evaluate the contribution of each CDR loop to EPCR binding. It is already evident from TCR chain swap experiments that the Vγ4 CDR3 loop is essential for EPCR binding (Willcox *et al*, 2012). Antibody-antigen binding usually involves all 6 CDR loops, while the G8 γδ TCR-T22 interaction is completely dependent on the CDR-3 loop of the δ chain (Adams *et al*, 2005). In addition, since the LES Vδ5 chain sequence contains many N-nucleotide additions, it is likely to result from a rare recombination event, suggesting it may be restricted to a minority of individuals. Dependence on such CDR loop sequences could explain why EPCR reactivity is a seemingly rare event. Since other EPCR reactive cells have not to date been identified, this may suggest all CDR loops may be required for binding. To address these issues, structural analysis of LES γδ TCR in complex with EPCR may give valuable insight into which CDR regions are important for ligand recognition.

CHAPTER 5 –

STRUCTURAL DETERMINATION OF SKINT-1

Chapter 5: Structural Determination of Skint-1

5.1 Introduction

Arguably one of the best studied γδ T cells subsets is the intraepithelial γδ TCR^+ cell subset found in the skin of mice. These cells have a characteristic dendritic morphology and are therefore known as dendritic epidermal T cells (DETC). DETCs are the first T cells to develop in the foetal mouse thymus, from embryonic day 13-17, and are subsequently only found in adult skin (Havran and Allison, 1990). Unlike αβ T cells and circulating γδ T cells, which utilise TCR V(D)J gene rearrangement and N-nucleotide addition by terminal deoxynucleotide transferase (TdT) to generate diversity in their TCRs, more than 90% of DETCs express a monomorphic Vγ5Vδ1 TCR, which is specifically recognised by the conformation-dependent 17D1 mAb. It is unclear whether the limited diversity of the canonical DETC TCR results from biased recombination of the gene segments used by DETC combined with the lack of TdT expression in the embryonic thymus during DETC development, or whether the lack of diversity is a result of positive selection. Early evidence favoured the biased recombination hypothesis, as the majority of thymocytes from mice expressing a Vγ5 gene segment containing a premature stop codon showed canonical rearrangements, even in the absence of TCR expression and hence positive selection. However, other experiments supported the idea of positive selection of DETC cells: in mice over-expressing TdT throughout embryonic development, more than 50% of in-frame TCR joins were canonical (Zhang *et al*, 1995). Further evidence for positive selection was obtained by disruption of the Vγ5 gene segment, which led to a replacement Vγ1 chain being utilised. Replacement Vγ1Vδ1 cells were still recognised by the 17D1 antibody, suggesting that the selection of a particular TCR conformation during their maturation in the thymus is important (Mallick-Wood *et al*, 1998).

DETCs, like the majority of γδ T cells, do not recognise classical peptide-MHC complexes. Instead, they appear to be activated by signals from adjacent damaged or stressed keratinocytes (Havran *et al*, 1991). These signals can include known ligands for the activating receptor NKG2D expressed by DETCs, as well as unidentified signals through the TCR. Once activated, whether through the TCR or NKG2D, DETCs are involved in regulating inflammation, modulating downstream immune responses, and maintaining epithelial integrity. Mice deficient in DETC display spontaneous ear inflammation that can be rescued by adoptive transfer of foetal $V\gamma5^+$ cells (Girardi *et al,* 2002). DETCs are also crucial in wound healing: mice lacking DETCs are slower to heal full-thickness skin wounds than wild type mice (Jameson *et al*, 2002). Following wounding, DETC near the wound lose their dendritic morphology, and produce growth factors such as FGF-7 and FGF-10 to encourage the regulated proliferation of keratinocytes around the wound site (Boismenu *et al,* 1994). Activated DETC also make chemokines that include MIP-1a/MIP-1b/RANTES, which attract monocytes and lymphotactin to T cells at the site of the wound (Boismenu *et al,* 1996), as well as pro-inflammatory cytokines IFN-γ and TNFα (Jameson *et al,* 2002). Other forms of keratinocyte damage, such as that generated *in vivo* or *in vitro* by exposure to contact sensitizers, also induce DETC proliferation *in vitro* (Huber *et al*, 1995). This has been shown to be mediated by the γδ TCR because anti-CD3 mAb effectively blocked activation of the DETC. Further studies using a doxycycline-inducible Rae-1 transgenic system examined the ability of DETC to organise downstream immune responses in the skin following activation through NKG2D but in the absence of an obvious TCR signal: induced Rae-1 caused DETC rounding, changes in the skin Langerhans cells, and recruitment of unconventional αβ T cells (Strid *et al*, 2008). However, recent intravital microscopy studies have suggested that even in the absence of stress stimuli, the γδ TCR of DETC is preferentially localised in contacts to

keratinocytes and provides steady state signalling to the DETC, suggestive of interaction with constitutively expressed ligands on keratinocytes (Chodaczek *et al*, 2012).

DETCs also play a role in protection against epithelial tumourigenesis. TCR$\delta^{-/-}$ mice showed increased susceptibility to several cutaneous chemical carcinogenesis regimes in the FVB/N strain (Girardi *et al*, 2001). This effect was shown to be specifically due to the lack of DETC by using mice deficient for both the Vγ5 and Vδ1 gene segments in the same models (Strid *et al*, 2008). Mice lacking either the Vγ5 or Vδ1 gene segments had DETCs which expressed "replacement" receptors that stained with 17D1 to some degree, and these cells protected against malignancy to the same degree as wild type DETC (Strid *et al*, 2008). DETC$^{-/-}$ mice showed increased incidence of papilloma formation compared to wild type mice, as well as increased progression from papilloma to carcinoma (Girardi *et al*, 2001; Strid *et al*, 2008), suggesting a role for DETC at very early stages of immunosurveillance. Ligands for the NKG2D receptor have been shown to be upregulated in murine skin within 24 hours of treatment with carcinogen (Girardi *et al*, 2001). In these models of tumourigenesis, it is likely DETCs are activated at least partly by signals through NKG2D, however γδ TCR ligands may also contribute, as supported by recent imaging studies (Chodaczek *et al*, 2012).

Despite extensive studies on DETCs, the ligand for the DETC TCR in the skin remains unclear, as does the identity of a positively selecting ligand in the thymus. However, a potential breakthrough came whilst using FVB/N mice from a new breeding colony at Taconic Farms (FVB_{tac}) as a control in the abovementioned tumour immunosurveillance studies (Girardi *et al,* 2001). Lewis, *et al* (2006) serendipitously found these FVB_{tac} mice had normal numbers of DETCs, but these cells expressed diverse TCRs and not the normal

monomorphic Vγ5Vδ1 TCR. These mice also displayed spontaneous cutaneous inflammation similar to $\delta^{-/-}$ mice. Backcrossing FVB_{tac} x FVB_{Jax} mice suggested an autosomal dominant gene was responsible for the selection of DETCs and genetic mapping ultimately identified a single substitution between FVB_{tac} and FVB_{Jax} mice, consisting of a mutation in the LOC384040 gene, which was characterised and named *Skint-1* (for selection and upkeep of intraepithelial T cells 1).

Skint-1 encodes a 364 amino acid protein which contains two Ig-like domains (one IgV-like and one IgC-like domain), three transmembrane domains and a short cytoplasmic tail at the C terminus (Figure 5.1). Skint-1 mRNA was only expressed in thymic epithelial cells, and also constitutively in keratinocytes (Boyden *et al,* 2008), consistent with a possible role in thymic selection of DETC as well as activation of mature DETC in the periphery. Expression of full-length skint1Jax rescues the development of mature Vγ5Vδ1 DETC in reaggregate thymic organ culture (RTOC), and FVB_{tac} mice which over-express transgenic wild type skint1 have normal numbers of DETCs in the skin (Boyden *et al,* 2008; Barbee *et al,* 2011). Therefore, it appears that Skint-1 expression in the thymus is essential for the selection of the DETC subset and homing to the epidermis, as the immature DETC which developed in FVB_{tac} mice were found to home to sites other than the skin and acquire a different functional phenotype (Turchinovich and Hayday, 2011). Over-expression of Skint-1 under the control of a β-actin promoter did not result in negative selection of DETC, or alterations in other γδ subsets (Barbee *et al,* 2011), suggesting DETCs do not undergo negative selection in the same manner as αβ T cells.

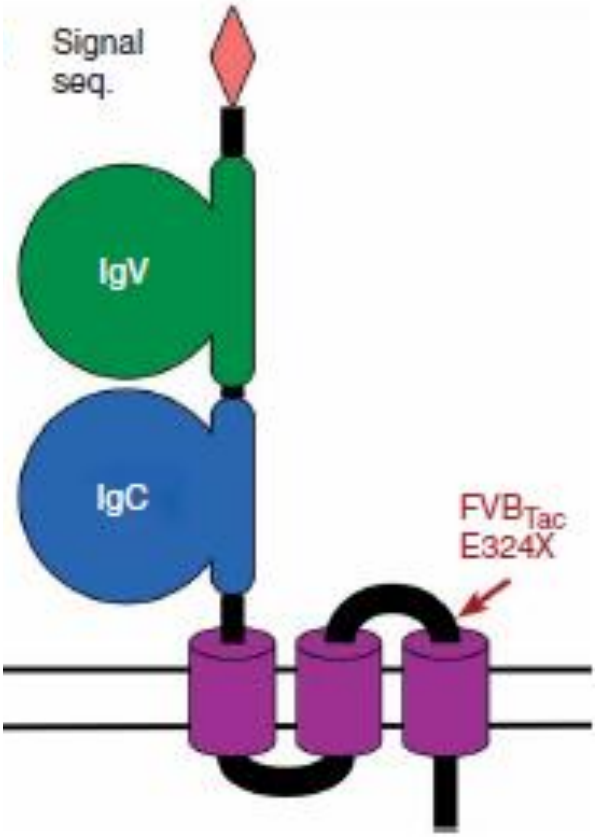


Figure 5.1 Schematic diagram of Skint-1 as expressed on the surface of thymic cells (Boyden *et al.*, 2008). Membrane-distal variable domain (IgV) is displayed in green and the membrane-proximal constant-like domain (IgC) shown in blue. Mutation identified in FVB/N mouse strain (generated by Taconic laboratories) that results in the selective deficiency for epidermal Vγ5Vδ1 T cells is highlighted.

In order to learn more about Skint-1, which clearly plays a key role in thymic selection of DETCs but is also a candidate ligand in the skin, I expressed Skint-1 and carried out structural studies on the recombinantly expressed protein. Recombinant production of soluble Skint-1 was undertaken in several expression systems, providing sufficient levels of protein, for downstream structural studies with X-ray Crystallography and NMR.

5.2 Production of Skint-1 ectodomain

5.2.1 Cloning and Expression

The mouse Skint-1 cDNA was kindly donated by the Hayday laboratory. The Skint-1 ectodomain construct consisting of the IgV and IgC regions (incorporating residues 1-224) was successfully amplified using Pfu Stratagene polymerase into a prokaryotic expression vector pET23a. Confirmation of the correct Skint-1 sequence cloning was verified using the Big Dye terminator kit from Functional Genomics. This expression vector was subsequently transformed into BL21 (DE3)pLYS *E. coli cells* and test expression for the Skint-1 IgV-IgC construct (in 5ml cultures) revealed a distinct band induced at ~31 kDa (data not shown). This highlighted that the Skint-1 protein was migrating non-ideally on the SDS-PAGE gel given that the predicted sequence derived molecular weight of the Skint-1 ectodomain is ~25.5 kDa. To obtain high yields of recombinant Skint-1, initial expression conditions were replicated on a large scale to a volume of 2L. Consistent with other Ig-domain containing proteins, centrifugation of lysed cultures of Skint-1 resulted in a large pellet of insoluble inclusion body material. Inclusion bodies were purified as described previously (Willcox *et al*, 1999) and finally solubilised in a denaturing buffer containing 8M urea. Generally, ~120 mg/litre of recombinant inclusion protein were obtained. SDS–PAGE analysis of purified inclusion body

pellets confirmed the presence of Skint-1 (Figure 5.2) at relatively high purity and sufficient levels for renaturation.

5.2.2 *In vitro* refolding of Skint-1 IgV-IgC

Initial attempts for renaturation of the Skint-1 ectodomain consisted of slowly removing the denaturant by dilution, a method that has proven successful for many other Ig-like domain-containing proteins (Allison *et al*, 2001 Garner *et al*, 2005). Renaturation was carried out in the presence of glutathione redox components to allow formation/breakage of disulphide bonds. In addition, an alkaline pH was maintained to promote formation of the active glutathione thiolate anion and hence disulphide exchange. Furthermore, L-arginine, an additive widely used to enhance *in vitro* protein refolding (Choi *et al*, 2012), was incorporated into the refolding buffer, to suppress protein aggregation. To reduce concentration-dependent aggregation effects, 60 mgs of Skint-1 IgV-IgC inclusion bodies were added in pulses to the refolding mixture over several days. Despite these provisions, the size exclusion profile of the refolded mixture revealed no prominent peak suggesting that the Skint-1 IgV-IgC protein was refractory to refolding (data not shown).

Further attempts at refolding the Skint-1 ectodomain comprised of supplementing the refold mixture with a denaturant, a strategy that has proven successful for renaturing TCRs (Willcox *et al*, 1999). The denaturant, in this case 5M urea, was included in the renaturation buffer for several days to prevent protein aggregation and then dialysed out. This strategy allowed the Skint-1 ectodomain to fold into its native state, as demonstrated by the size exclusion profile following purification on the Superdex 200 column (Figure 5.3). However, the recovery yield of the Skint-1 IgV-IgC was extremely inefficient (0.008%), suggesting that the levels of

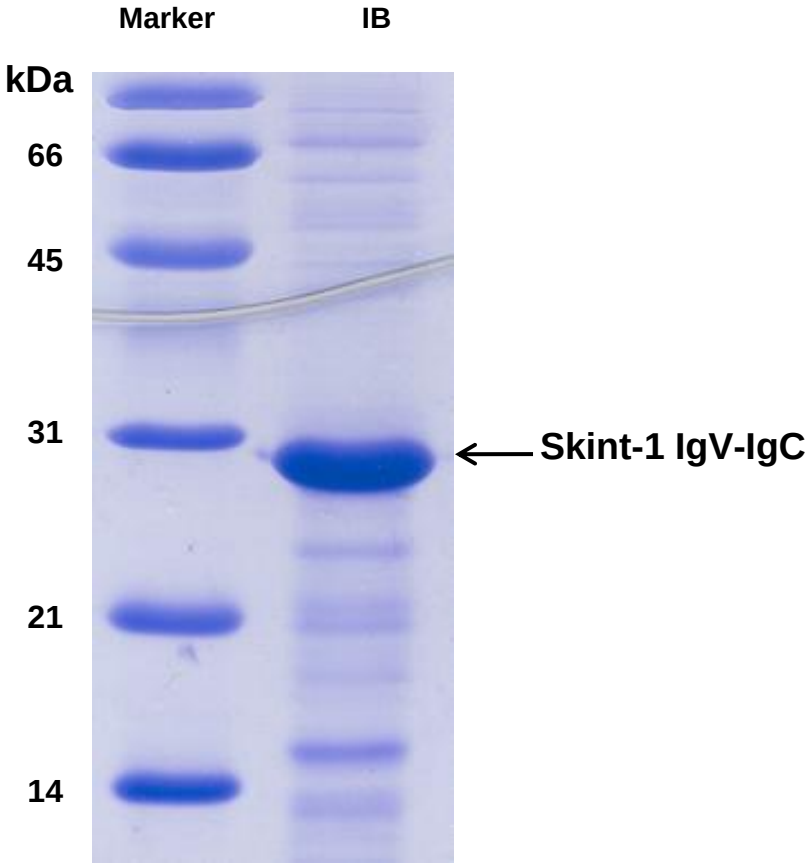


Figure 5.2 SDS-PAGE analysis of the Skint-1 ectodomain expressed as inclusion bodies in *E. coli*. Molecular weight standards are shown in the first lane and inclusion body preparations in the second lane. Band corresponding to Skint-1 IgV-IgC (black arrow; Mw 26 kDa) has migrated close to the 31 kDa standard. Full-length Skint-1 inclusion body preparations were relatively pure (>80%) and contained high levels of protein (120 mg/L).

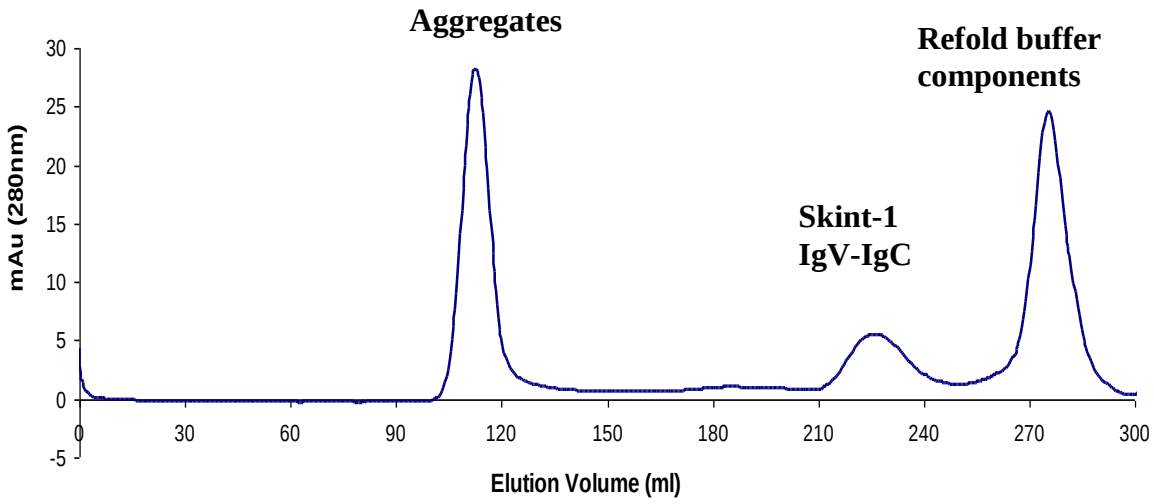


Figure 5.3 Elution profile for Skint-1 IgV-IgC ectodomain following purification by size exclusion chromatography using the Superdex 200 column. Skint-1 ectodomain elutes at 225 ml consistent with the Skint-1 ectodomain existing as a monomer in solution (26kDa). From a 60 mg Skint-1 pulsed refold a peak of 5 mAu is observed which is unfeasible for downstream structural studies.

correctly refolded protein was inadequate for downstream structural studies. One possibility for enhancing recovery yields consisted of scaling up the refold mixture and pulsing with increased concentrations of inclusion bodies. However, given that the levels of protein required for structural and binding studies are relatively high, this approach was deemed to be impractical, expensive and time-consuming.

5.3 Recombinant Production of Skint-1 IgV domain

5.3.1 Cloning, expression and inclusion body preparation

Given that the Skint-1 ectodomain was intransigent to refolding and therefore an unsuitable candidate for structural analyses, a new approach was employed which consisted of producing a single immunoglobulin variable (IgV) domain of Skint-1. Notably, a previous member of the Willcox group having encountered similar renaturation problems with the LILRB4 ectodomain (Domains 1 and 2), successfully refolded LILRB4 IgV-like domain in amounts sufficient for crystallisation trials (PhD Thesis, L.I. Garner, 2011). To generate the Skint-1 IgV construct, a stop codon (TAA) was introduced at the C-terminal end of the IgV domain (residue 119T) in the pET23a expression construct. The domain boundary for the Skint-1 IgV domain was determined using a combination of secondary structure predictions and structural homologues. Following verification of the Skint-1 IgV sequence by DNA sequencing, small-scale test expressions (in 5 ml cultures) were performed and analysed by SDS-PAGE. This revealed an over-expressed protein band at ~13.5 kDa in the induced sample relative to the un-induced cells (Figure 5.4a), consistent with the expression of the Skint-1 IgV domain. The single IgV domain was subsequently expressed as inclusion bodies on a large scale (in 2L cultures), with estimated yields of 136 mg/L based on the Bradford assay. Analysis of the Skint-1 IgV inclusion body preparations by SDS-PAGE demonstrated high quality protein for

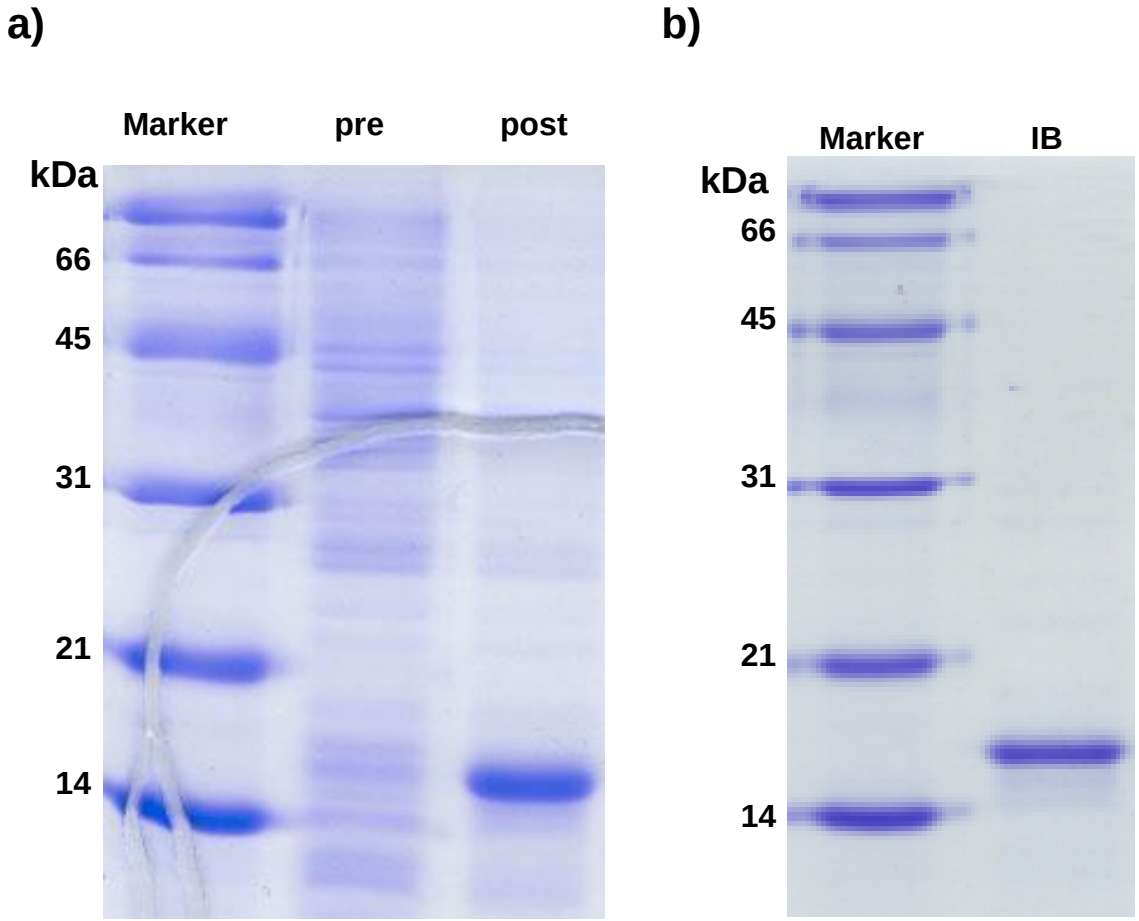


Figure 5.4 SDS–PAGE analysis of Skint-1 IgV domain test expression and inclusion body preparation in BL21 *E. coli* cells. a) Test expression analysis reveals Skint-1 IgV over-expression in post-induction sample (Lane 3) relative to pre-induced sample (Lane 2). b) Skint-1 IgV inclusion body material solubilised in 8M urea (IB). Molecular weight standards are shown in the first lane, masses indicated in kDa.

renaturation studies (Figure 5.4b). In addition, analysis of purified inclusion bodies using the SALDI-TOF mass spectrometer revealed a 13.5 kDa peak, consistent with the sequence derived molecular weight of the Skint-1 IgV domain (Figure 5.5).

5.3.2 Renaturation and purification of Skint-1 IgV

The Skint-1 IgV inclusion bodies were refolded using the oxidative *in vitro* dilution method (Willcox *et al.*, 1999). Confirmation of correctly refolded protein was assessed by size exclusion chromatography and SDS-PAGE. The size exclusion elution profile revealed a large peak (245 mAu) eluting at 235 ml in 20 mM Tris pH 8 and 50 mM NaCl (Figure 5.6), consistent with Skint-1 IgV existing as a monomeric form in solution. Importantly, the refold yield for the IgV domain was considerably improved relative to the IgV-IgC ectodomain (5 mAu). Moreover, SDS-PAGE analysis of the Skint-1 IgV peak fraction revealed sufficient levels of purity for subsequent crystallisation trials (Figure 5.7).

5.4 Structural studies of Skint-1 IgV domain using X-ray Crystallography

5.4.1 Crystallisation trials of Skint-1 IgV

Purified Skint-1 IgV was concentrated down to 5-15 mg/ml and subjected to crystallisation trials using the hanging drop vapour diffusion method with the Mosquito nanolitre crystallisation robot. Initial crystallisation trials were carried out using the commercial sparse matrix kits with the hanging drop vapour diffusion method. Using this approach two potential nanocrystal hits were observed in drops equilibrated against the PACT screen condition 1-14 (0.1M MIB buffer pH5, 25% w/v PEG 1500) and 1-41 (0.1M MMT buffer pH8, 25% w/v PEG 1500) (Figure 5.8 a-b). However, these hits were exceedingly difficult to reproduce and subsequent optimisation trials involving adjustments in the pH and variations in the protein

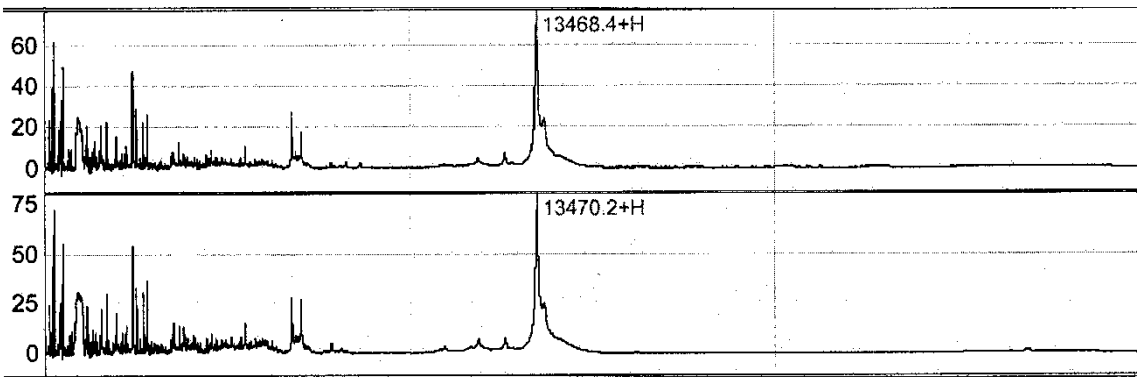


Figure 5.5 Mass spectrometry analysis of Skint-1 IgV inclusion bodies. Sample was analysed by SALDI TOF mass spectrometer. Top and bottom spectrograms correspond to replicate samples of Skint-1 IgV with the size of the protein demonstrated to be ~13.5 kDa.

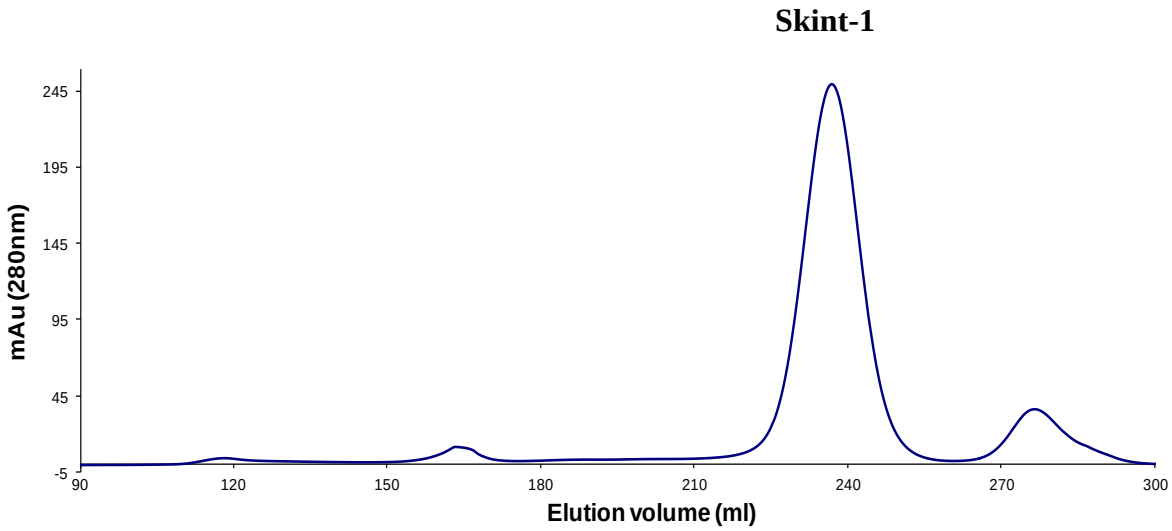


Figure 5.6 Elution profile for Skint-1 IgV following purification by size exclusion chromatography using the Superdex 200 column. Skint-1 IgV domain elutes at 235 ml consistent with existing as a monomer in solution (14kDa). From a 30 mg pulsed refold a peak of 245 mAu is observed which was sufficient for downstream structural studies.

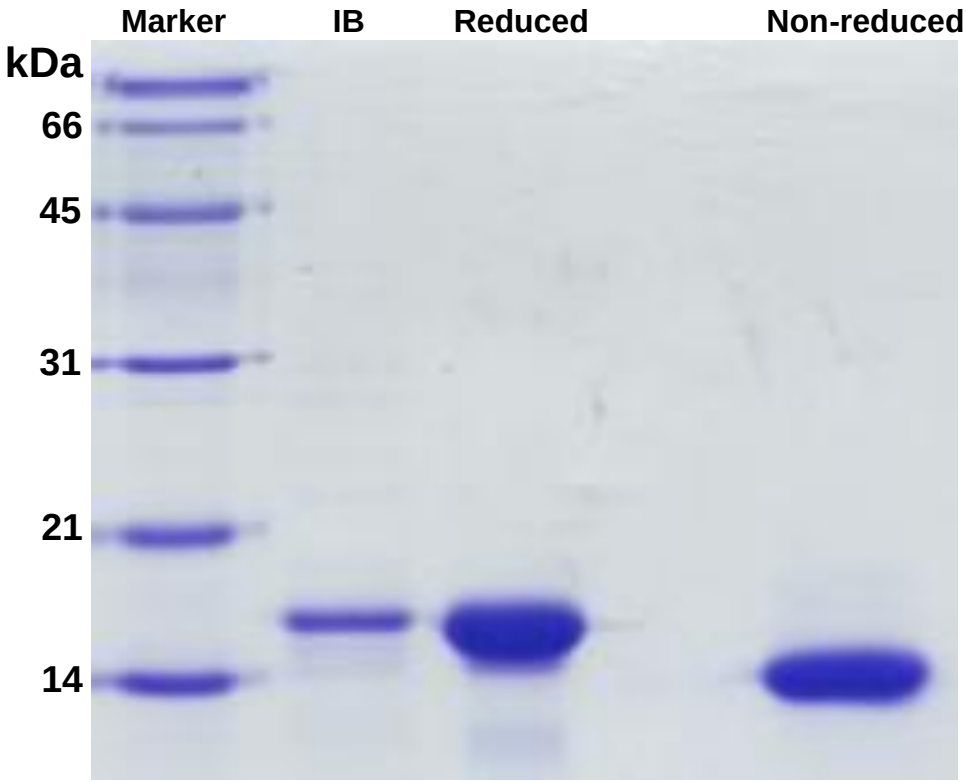
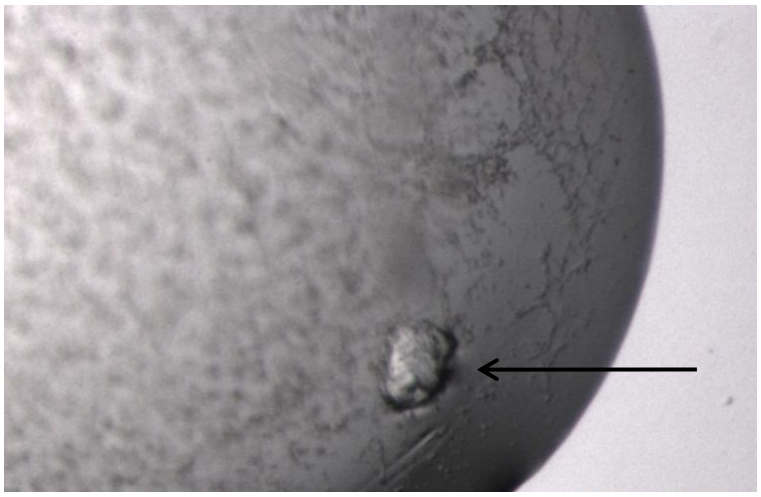


Figure 5.7 SDS-PAGE analysis of refolded Skint-1 IgV domain. Skint-1 IgV was renatured and purified by size exclusion chromatography using the Superdex 200 column. Molecular weight standards are shown in the first lane, masses indicated in kDa. Lanes 2, 3 and 4 show inclusion body (IB), reduced and non-reduced samples of Skint-1 IgV, respectively.

a)



b)

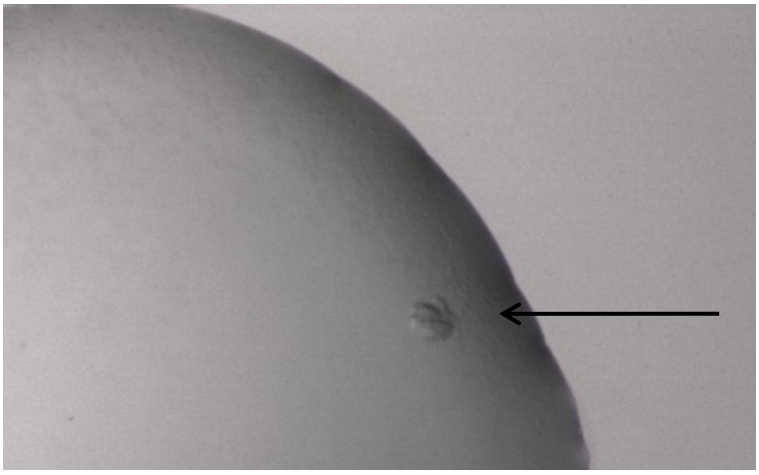


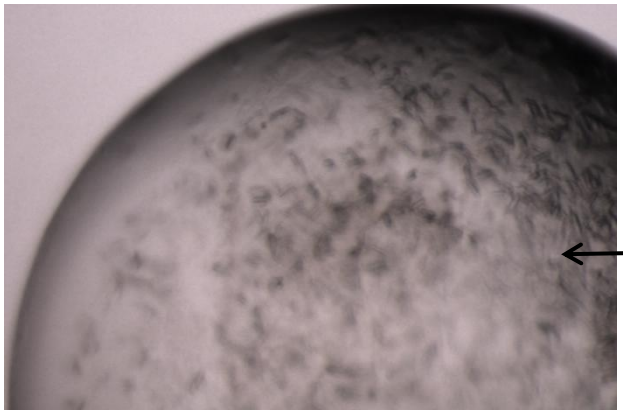
Figure 5.8 Micro-crystals of Skint-1 IgV grown using the hanging drop vapour diffusion method. a) The drop was set up by mixing Skint-1 IgV (10mg/ml in 50mM NaCl/20mM Tris pH8) with an equal volume of reservoir buffer (PACT screen condition number 1-14, 0.1M MIB pH5,25% w/v PEG 1500). b) The drop was set up by mixing Skint-1 IgV (10mg/ml in buffer) with an equal volume of reservoir solution (PACT screen condition number P1-41, 0.1M MMT pH8, 25% w/v PEG 1500). In each case the crystals (black arrow) appeared after 7 days at room temperature.

and precipitant concentration led to no further improvement in crystal size or morphology. As a result, the Skint-1 IgV protein was subjected to additional crystallisation screening trials. Based on these extended trials, preliminary crystallisation hits were identified in 2M ammonium sulphate and 0.1M Tris pH 8 (Figure 5.9a). Strikingly, these trials when repeated in the absence of protein revealed no crystals thereby considerably reducing the possibility that these crystals were derived from salt. However, this particular crystallisation condition proved to be equally challenging for generating diffraction-grade Skint-1 IgV crystals. Hence, further optimisation trials were performed with the detergent screen to improve the size and morphology of the crystals. The most favourable results were gained from the addition of BigCHAPS (1.4mM) to the existing condition (2M ammonium sulphate and 0.1M Tris pH 8) at room temperature and 4°C (Figure 5.9b). Although, this condition significantly improved the size of these crystals, their morphology was unsuitable for X-ray diffraction experiments. Hence, further optimisation trials with the additive screen were carried out to improve the morphology of the Skint-1 IgV crystals. Using this approach, PEG 400 was identified as a highly promising additive that promoted formation of large diamond shaped crystals (Figure 5.10a). The final optimised condition for reproducibly generating Skint-1 IgV crystals for crystallographic analysis consisted of 2M ammonium sulphate, 0.1M Tris pH 8, 1.4 mM BigCHAPS and 5% PEG 400. These crystals were grown manually on a large scale (2µl + 2µl drops), favouring growth at 4°C and appearance of crystals after 1-2 weeks (Figure 5.10b).

5.4.2 X-ray diffraction studies of Skint-1 IgV

Single Skint-1 IgV crystals were soaked in mother liquor supplemented with increasing concentrations of either ethylene glycol (5-23%) or glycerol (5-20%) and flash cooled in liquid nitrogen. X-ray diffraction studies were carried out using an ‘in house’ rotating anode

a)



b)

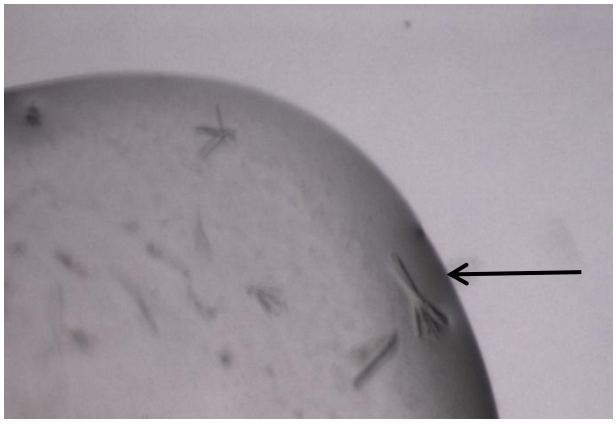


Figure 5.9 Generation and optimisation of Skint-1 IgV crystals. a) Micro-crystals of Skint-1 IgV were identified using the ProPlex Screen. The drop was set up by mixing Skint-1 IgV (10mg/ml in 50mM NaCl/20mM Tris pH8 buffer) with an equal volume of reservoir solution (condition number 76, 2M NH_2SO_4/0.1M Tris pH8). b) Optimisation of Skint-1 IgV crystals were carried out using the Detergent screen. The best crystals were grown by incorporating 1.4mM BigCHAPS to the existing crystallisation condition (2M NH_2SO_4/0.1M Tris pH8). The crystals (black arrow) were grown using the hanging drop vapour diffusion method and appeared after 7 days at room temperature and 4°C.

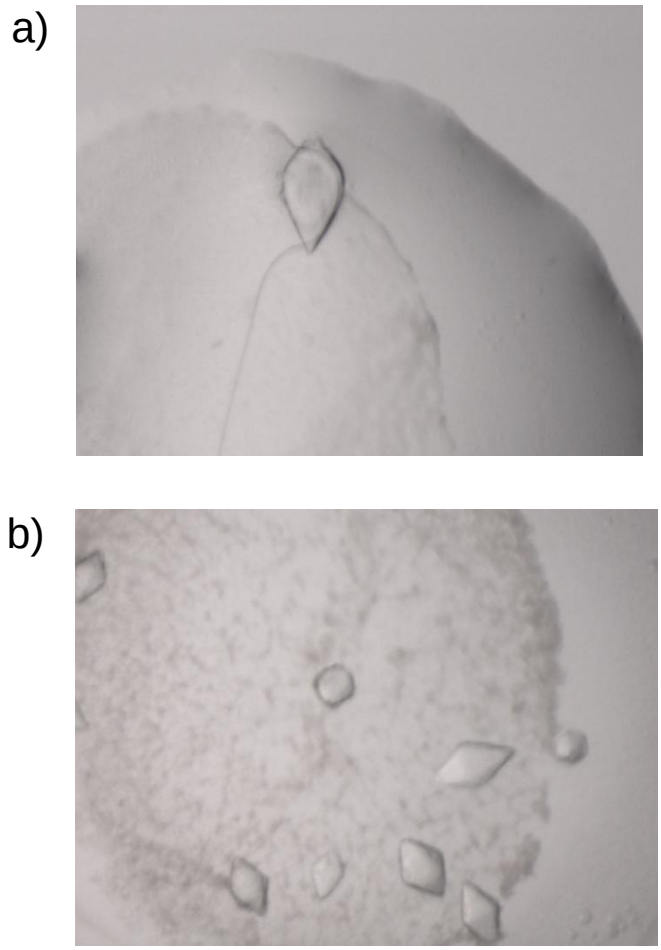


Figure 5.10 Generation of diffraction-grade Skint-1 IgV crystals. a) Optimisation of Skint-1 IgV crystals were carried out using the additive screen. Diamond shaped crystals of Skint-1 IgV were grown by incorporating 5% PEG 400 to the existing crystallisation condition (2M NH_2SO_4, 0.1M Tris pH8 and 1.4 mM BigCHAPS). The drop was set up by mixing 300 nl of Skint-1 IgV (10mg/ml in 50mM NaCl/20mM Tris pH8 buffer) with an equal volume of crystallisation solution. b) Large-scale Skint-1 IgV crystals were grown in 2µl + 2µl drops using crystallisation conditions described in the left panel. The crystals were grown using the hanging drop vapour diffusion method and appeared after 5 days at room temperature and 4°C.

Rigaku X-ray generator and a Saturn CCD detector. All X-ray diffraction experiments were performed at 100K. An initial Skint-1 IgV crystal pre-soaked in 23% ethylene glycol for 1 hour diffracted X-rays to 4.5Å resolution (data not shown). However, crystals soaked in 20% glycerol for longer intervals (up to 30-40 minutes) diffracted to a higher resolution (Figure 5.11).

The best Skint-1 IgV crystal diffracted to a resolution of 4Å and crystallised in the trigonal/hexagonal space group with unit cell parameters a=73.8=b=73.8Å, c=140.7 Å, α=β=90°, γ=120°. The data collection strategy was calculated using the CrystalClear software. A total of 720 images each covering an oscillation angle of 0.5°, were collected with an exposure time of 15s per frame and a crystal to detector distance of 100 mm. Intensity data from the Skint-1 IgV crystal were scaled and merged using programs of the XDS suite. The relevant data processing statistics for Skint-1 IgV are listed in Table 5.1.

Resolution Limits	20-4Å
Unit Cell Dimensions	a=b=73.8 Å c=140.7Å α=β=90° γ=120°
Space Group	P622
Total Observations	43923 (2874)
Unique Observations	2038 (127)
Multiplicity	21.5 (22.6)
Completeness (%)	98.4 (97.7)
R_{merge} (%)	9.5 (76.8)
I/σI	31.1 (6.1)
Redundancy	21.6 (22.6)

Table 5.1. Data processing statistics for Skint-1 IgV*. X-ray data were processed using programs of the XDS suite. Numbers in parentheses correspond to statistics in the highest resolution shell. The data processing statistics were acceptable for subsequent structural determination analysis.*

In order to overcome space group ambiguity (trigonal/hexagonal), the data were processed in the lower symmetry trigonal space group, P3, and input into POINTLESS (Winn *et al*, 2011).

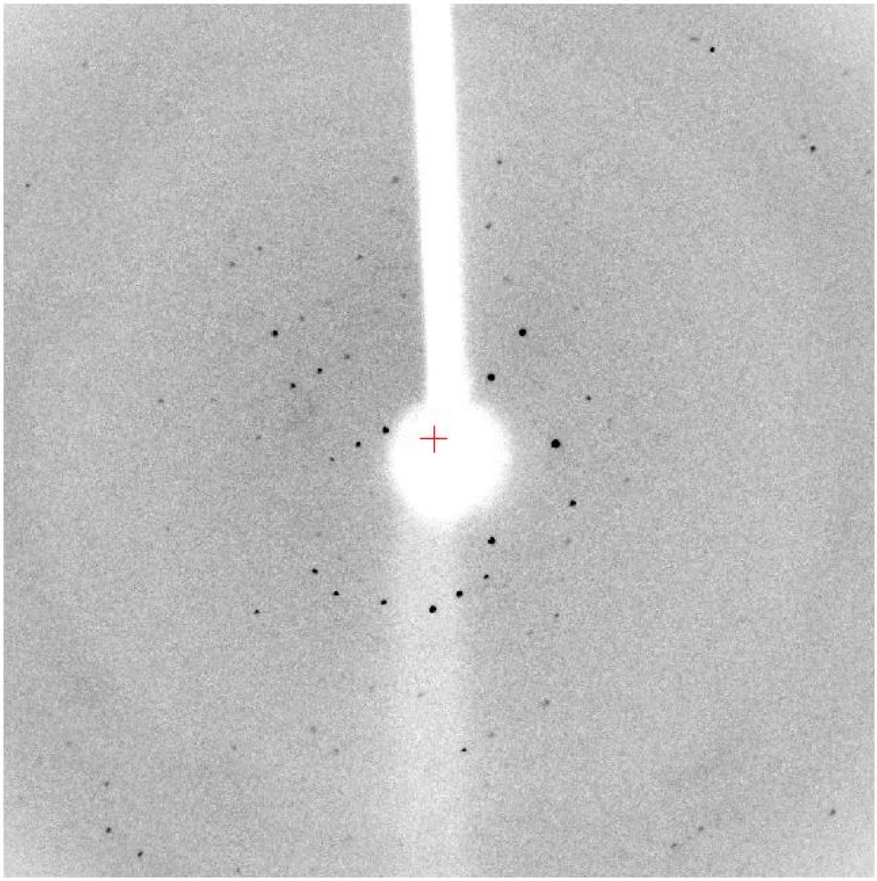


Figure 5.11 A 0.5° oscillation image of Skint-1 IgV crystal collected at 100K with a Saturn CCD detector using a rotating anode Rigaku X-ray generator. The crystal was soaked in mother liquor supplemented with 23% glycerol. The crystal diffracted X-rays to a resolution of 4Å and crystallises in the hexagonal space group P622.

The output from POINTLESS indicated that the best possible space group solution was P622. However, unequivocal space group determination could only be made following a correct translation function solution during molecular replacement.

5.4.3 Determination of phase information with Molecular Replacement

Assuming the Skint-1 IgV crystals have two molecules per asymmetric unit, a solvent content of 40% was calculated (Kantardjieff KA and Rupp B, 2003). Following the submission of the Skint-1 IgV sequence to BLAST, a putative structural homologue, murine myelin oligodendrocyte (MOG) (PDB code 1PY9), was identified. The homologue shared 47% sequence identity to Skint-1 IgV, and unsurprisingly, the overall fold of MOG conformed to the classical IgV type fold. However, no significant rotation function solutions were found using MOG as the phasing model. The lack of success in determining phase information with molecular replacement may have been due to the low quality of the X-ray diffraction data or that the phasing model possessed regions that were structurally different to Skint-1. It was theoretically possible to calculate phase information by alternative methods such as MAD or MIR. However, both techniques were deemed unfeasible due to poorly diffracting Skint-1 IgV crystals. Based on these considerations, it prompted the idea of determining the solution structure of Skint-1 IgV using nuclear magnetic resonance (NMR).

5.5 Structural studies of Skint-1 IgV domain using NMR

5.5.1 Production of Labelled Skint-1 IgV for NMR analysis

The X-ray diffraction studies of Skint-1 IgV crystals yielded low resolution data that prevented structure determination using the molecular replacement technique. Hence, to elucidate the detailed structural features of Skint-1 IgV domain, attempts were made to

determine the solution structure of this protein with NMR. Experiments and analysis were performed in collaboration with Dr Timothy J Knowles (University of Birmingham). To assess the potential of utilising the *E. coli* derived Skint-1 IgV material for NMR analysis, the sample was initially buffer exchanged into 0.1M MES pH 6.5 to reduce amide exchange effects that occur at higher pH levels, and was subjected to one dimensional (1D) 1H-NMR spectroscopy. Buffer exchange was required as the higher exchange rate with solvent at basic pH leads to a significant loss of signal. Analysis of the resulting 1D 1H-NMR spectrum (Figure 5.12) demonstrated that the Skint-1 protein was folded and provided a good signal-to-noise ratio. This was demonstrated by good dispersion of peaks, especially in the methyl region with peaks clearly observed down to -1ppm as well as in the amide and aromatic region with signal dispersed from 6–10 ppm. Based on these considerations, the Skint-1 IgV domain represented a highly promising candidate for structural determination by NMR spectroscopy.

In order to conduct further analysis, isotope labelling of Skint-1 IgV was required. Briefly, single labelled (^{15}N) Skint-1 IgV was grown in M9 media, renatured by dilution refolding and purified by size exclusion chromatography as described in Chapter 2. Based on the size exclusion elution profile, ^{15}N Skint-1 IgV eluted as a monomer at approximately 240 ml similar to the unlabelled protein (Figure 5.13a). Similar results were obtained for the double labelled ^{13}C/^{15}N Skint-1 IgV protein (Figure 5.13b). Notably, the recovery yield for the labelled proteins was as efficient as the unlabelled sample. SDS-PAGE analysis of the labelled recombinant proteins demonstrated that the sample was homogeneous and thus ideal for NMR analysis (Figure 5.14a-b).

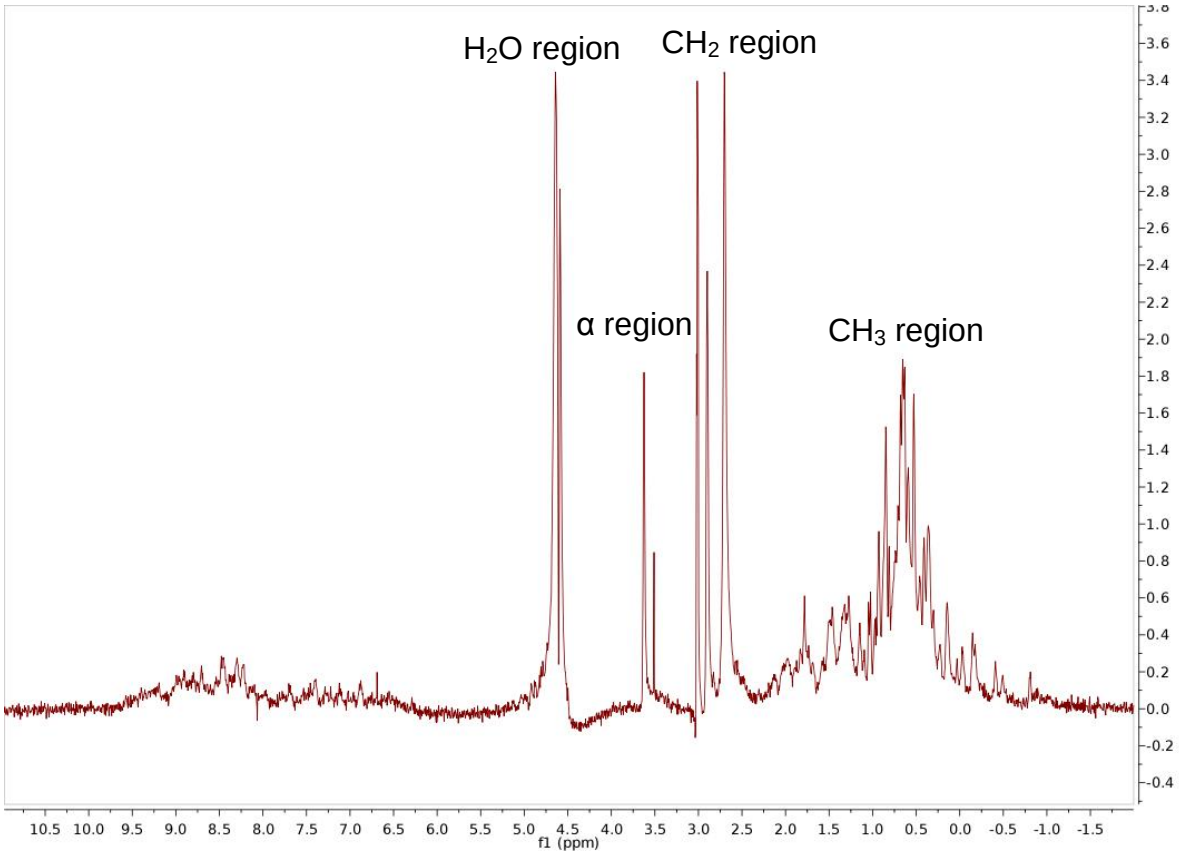


Figure 5.12 One dimensional ^{1}H spectrum profile of Skint-1 IgV. The Spectrum profile indicates that Skint-1 IgV domain data (50mM NaCl/20mM MES pH6.5) is amenable to NMR analysis. Chemical shift peaks are observed down to -1ppm in the proton scale (x axis), that are sharp and dispersed peaks and clearly visible in the CH_3 region, down to 9.5ppm.

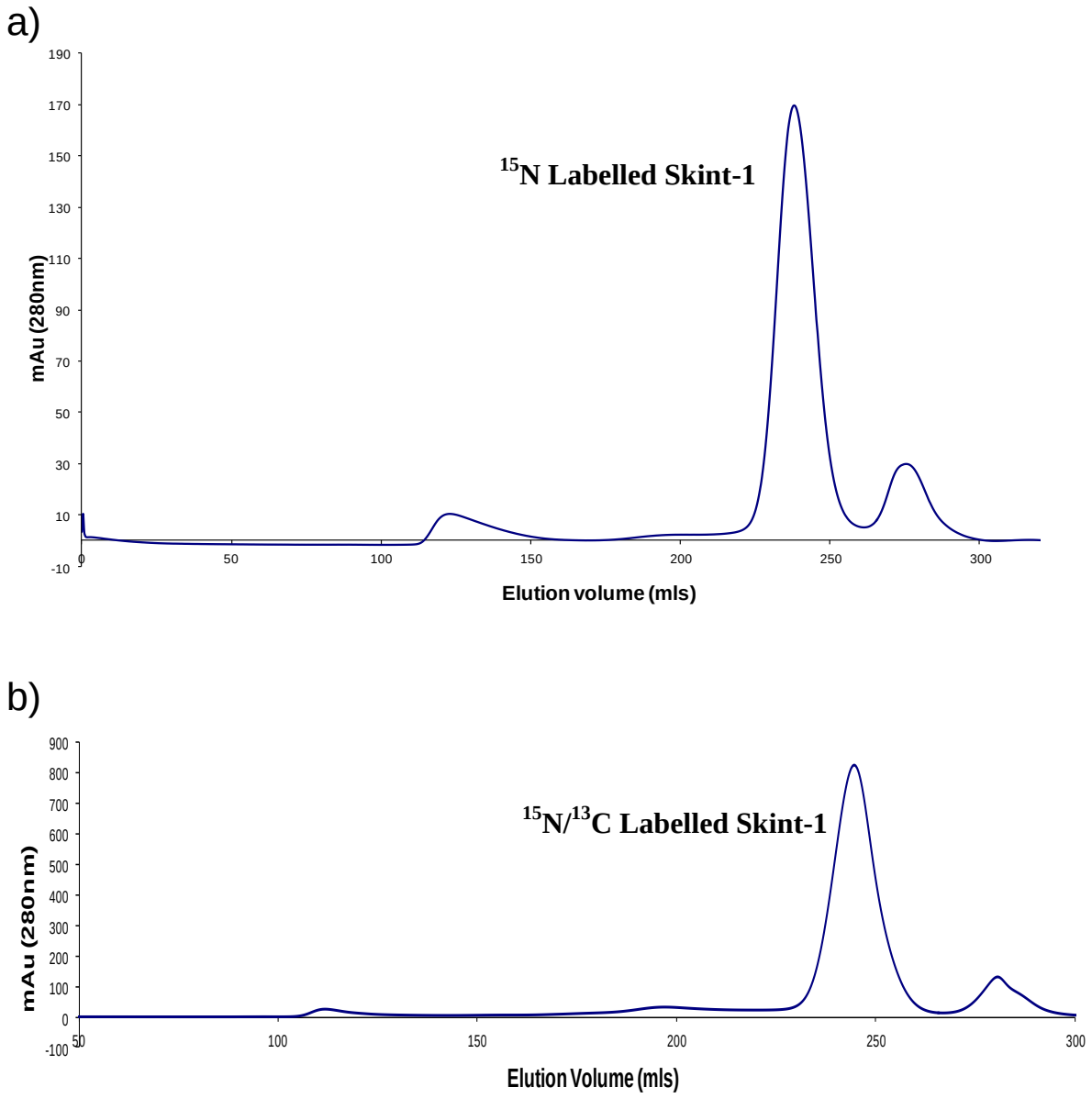


Figure 5.13 a) Elution profile for ^{15}N labelled Skint-1 IgV following purification by size exclusion chromatography using the Superdex 200 column. Single labelled Skint-1 IgV elutes at 235 ml consistent with existing as a monomer in solution (~14kDa). From a 60 mg pulsed refold a peak of 173 mAu is observed. b) Elution profile for $^{15}N/^{13}C$ labelled Skint-1 IgV following purification by size exclusion chromatography using the Superdex 200 column. Double labelled Skint-1 IgV elutes at 235 ml consistent with existing as a monomer in solution (~14kDa). From a 60 mg pulsed refold a peak of 815 mAu is observed.

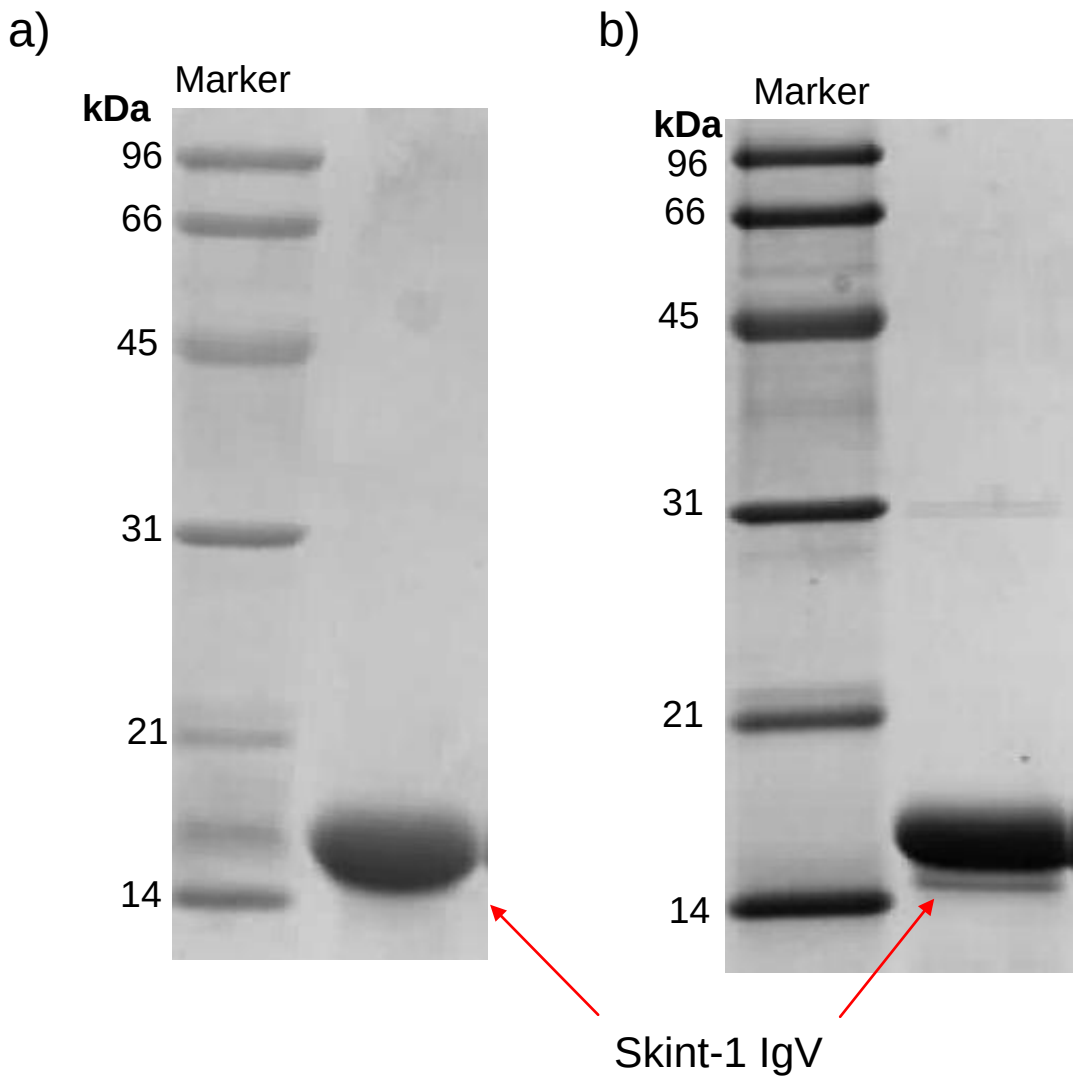


Figure 5.14 SDS-PAGE analysis of refolded and size exclusion purified labelled Skint-1 IgV. a) ^{15}N labelled protein. b) $^{15}N/^{13}C$ labelled protein in 50mM NaCl and 20mM MES pH 6.5

5.5.2 HSQC analysis

The heteronuclear single quantum coherence (HSQC) experiment provides a heteronuclear two dimensional spectrum by detecting a correlation between protons and attached hetero nuclear atoms either $^{13}C/^{15}N$. For the purpose of this thesis only the ^{1}H ^{15}N HSQC was utilised. This experiment highlights the one bond correlation between protons and directly attached nitrogen's, which in a protein are the amides present in the backbone and amines present in the side chains of Asparagine and Glutamine predominantly. The proton dimension is described in the x-axis and the nitrogen dimension is highlighted in the y-axis (Figure 5.15)

^{15}N Skint-1 IgV was analysed at a concentration of 19 mg/ml in 20 mM MES pH 6.5 and 50 mM NaCl. The HSQC spectrum for ^{15}N Skint-1 IgV domain demonstrated that the protein was stable and gave an ideal, well dispersed spectrum showing the protein was folded (Figure 5.15). The HSQC data set shows 93 clearly defined amide peaks out of a potential 114 resonances which could be observed (excluding 5 unobservable proline residues), suggesting 81% of the protein could be clearly resolved. Hence, the Skint-1 IgV domain was an ideal candidate for assignment and structure calculation using NMR spectroscopy.

5.5.3 Backbone assignments

To interpret the data recovered from NMR experiments, it is important to assign the resonances. In proteins, the most expedient method is to assign the backbone. This can be achieved by linking the atoms to their sequential neighbours, correlating via the amides and in essence walking along the protein backbone. The individual experiments listed allowed this walk along the protein. Backbone assignments were derived from BEST versions of ^{1}H-^{15}N-HSQC, HNCACB, HN(CO)CACB, HNCA, HN(CO)CA, HNCO and HN(CA)CO

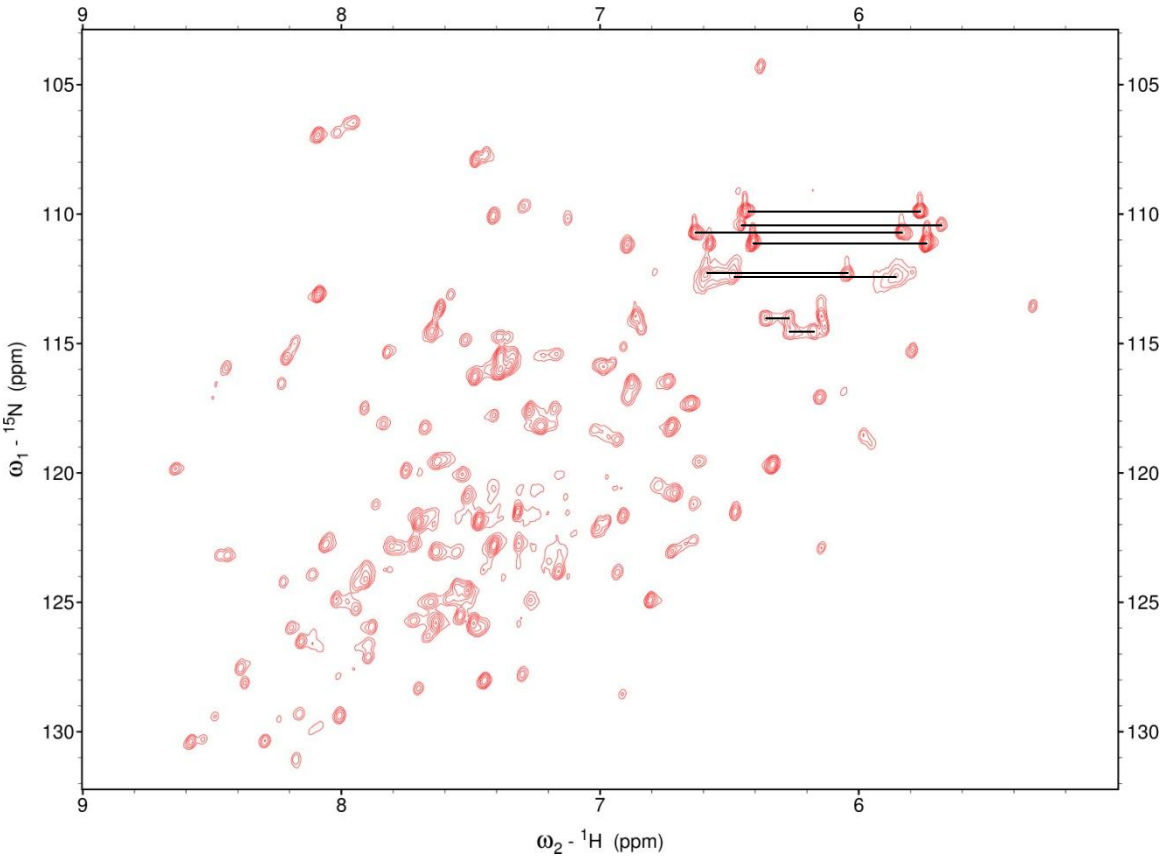


Figure 5.15 Two dimensional HSQC spectrum of ^{15}N labelled Skint-1 IgV. The Spectrum was recorded at 303 K on a Varian Inova 600 MHz NMR spectrometer equipped with triple resonance cryogenic probe and z-axis pulse field gradients. Skint-1 IgV was at 19mg/ml (50mM NaCl/20mM MES pH6.5). This spectrum profile indicates the protein is folded and has dispersion of signals from 5 to 9 ppm, with little signal overlap of each peak which represents an amide in the protein backbone. Peaks joined by lines correspond to amine side chain peaks of Asparagine and Glutamine.

experiments (Muhandiram DR 1994) (Figure 5.16). The description of the HNCA experiment correlates the proton and nitrogen of the backbone amide of residue i with its carbon α resonance and that for i-1 for a particular residue (Figure 5.16). This provided both intra-residue and inter-residue data and therefore enabled a walk along the backbone of Skint-1 IgV. The HN(CO)CA experiment is similar to the HNCA experiment but this correlates only the HN and N resonances for residue i and that of the Cα residue of i-1 (Figure 5.17), so distinguishing between the two resonances observed in the carbon dimension of the HNCA (Figure 5.17). By matching the i-1 Cα resonance associated with one HN, with the i Cα resonance associated with another HN, it was possible to identify the subsequent amino acid in the backbone chain and thus move along it. The remaining experiments were undertaken to overcome the difficulty of overlapping Cα signals in more complex larger proteins by utilising Cβ and CO inter-residue links as well, the HNCACB and HN(CO)CACB experiments provide the Cβ chemical shift, whilst the HNCO and HN(CA)CO provides the carbonyl.(Figure 5.17)

The ^{1}H ^{15}N HSQC of the Skint-1 IgV domain is shown in Figure 5.18. In collaboration with Dr. Timothy J Knowles, backbone assignments were completed for all non proline ^{1}H, ^{13}C and ^{15}N backbone resonances except for the amides of M1, H37, G57, S66, and those in the regions 44 to 52 and 69 to 71, presumably due to the dynamic nature of these residues (Table 5.2).

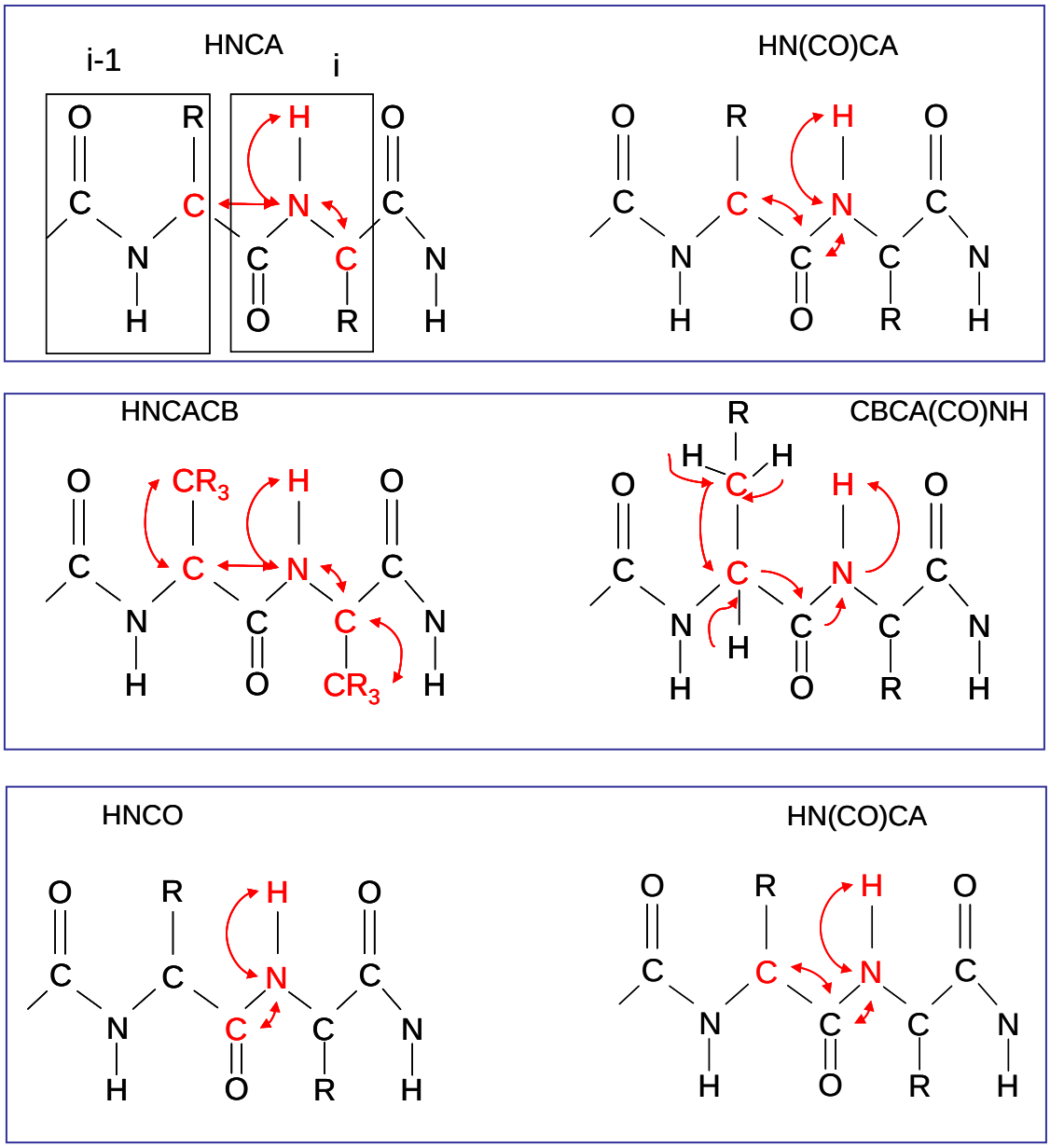


Figure 5.16 NMR experiments that were utilised to assign the backbone structure of Skint-1 IgV. Highlighted by blue boxes are the pairs of experiments that facilitated the construction of the backbone residues, atoms in red indicate those residues for which chemical shift data was recorded. Shown by the HNCA experiment in black boxes are the intra (i) and inter (i-1) residue signals that allow for pairing of residue assignments. The arrows represent the direction of magnetic transfer.

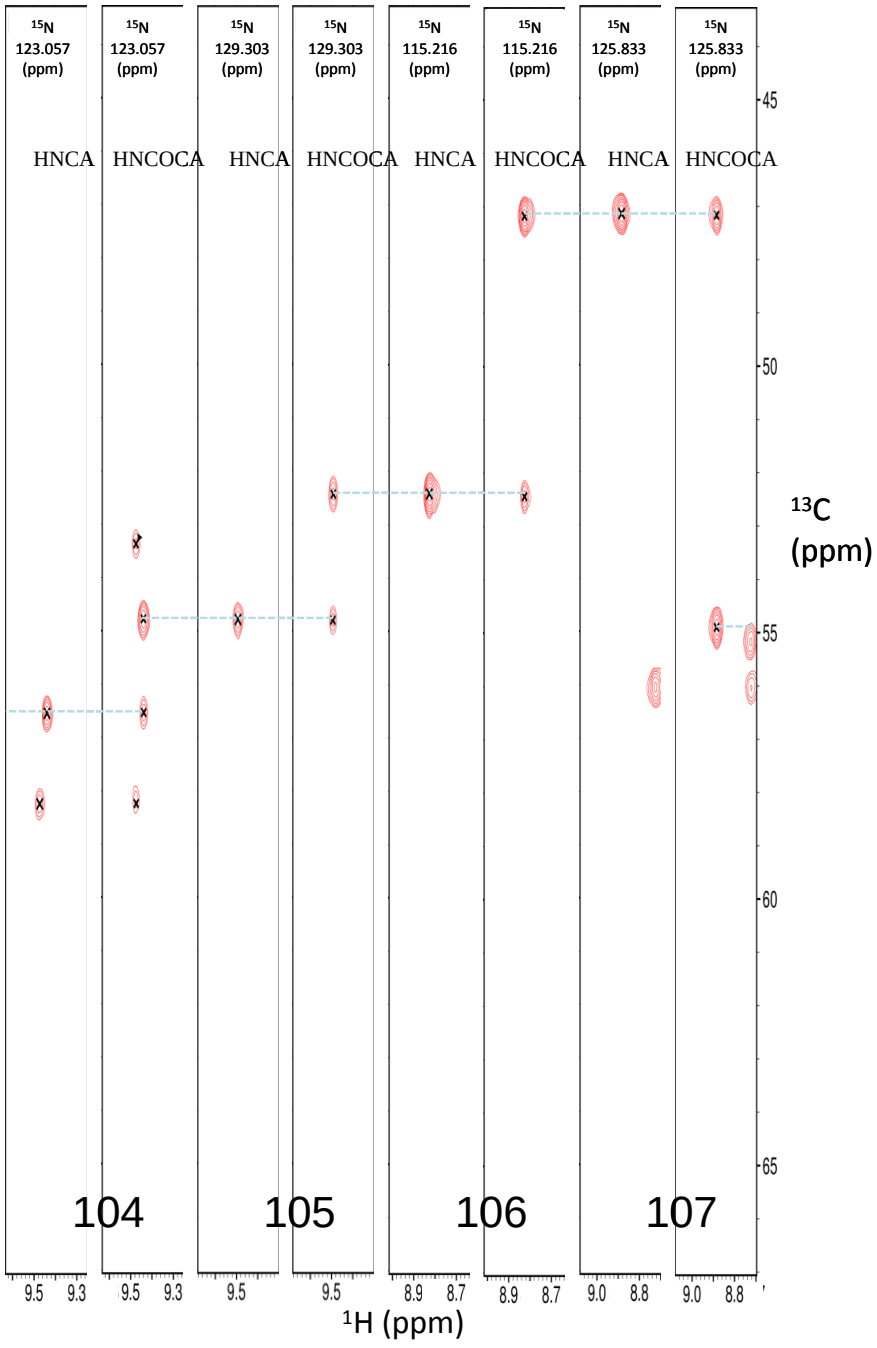


Figure 5.17 Strip plots from residues 104 to 107 highlighting the proton and carbon planes taken from the HNCA and HN(CO)CA. Shown in dashed lines are resonances associated with each residue and highlight the interaction between residue i and i-1. The amide proton chemical shift is displayed in the x axis and the Cα chemical shift is in the y axis. The nitrogen chemical shifts are taken from variable z axis's as displayed at the upper section of the strip plots. The dotted lines represent the inter-chain walk through of the protein backbone assignments.

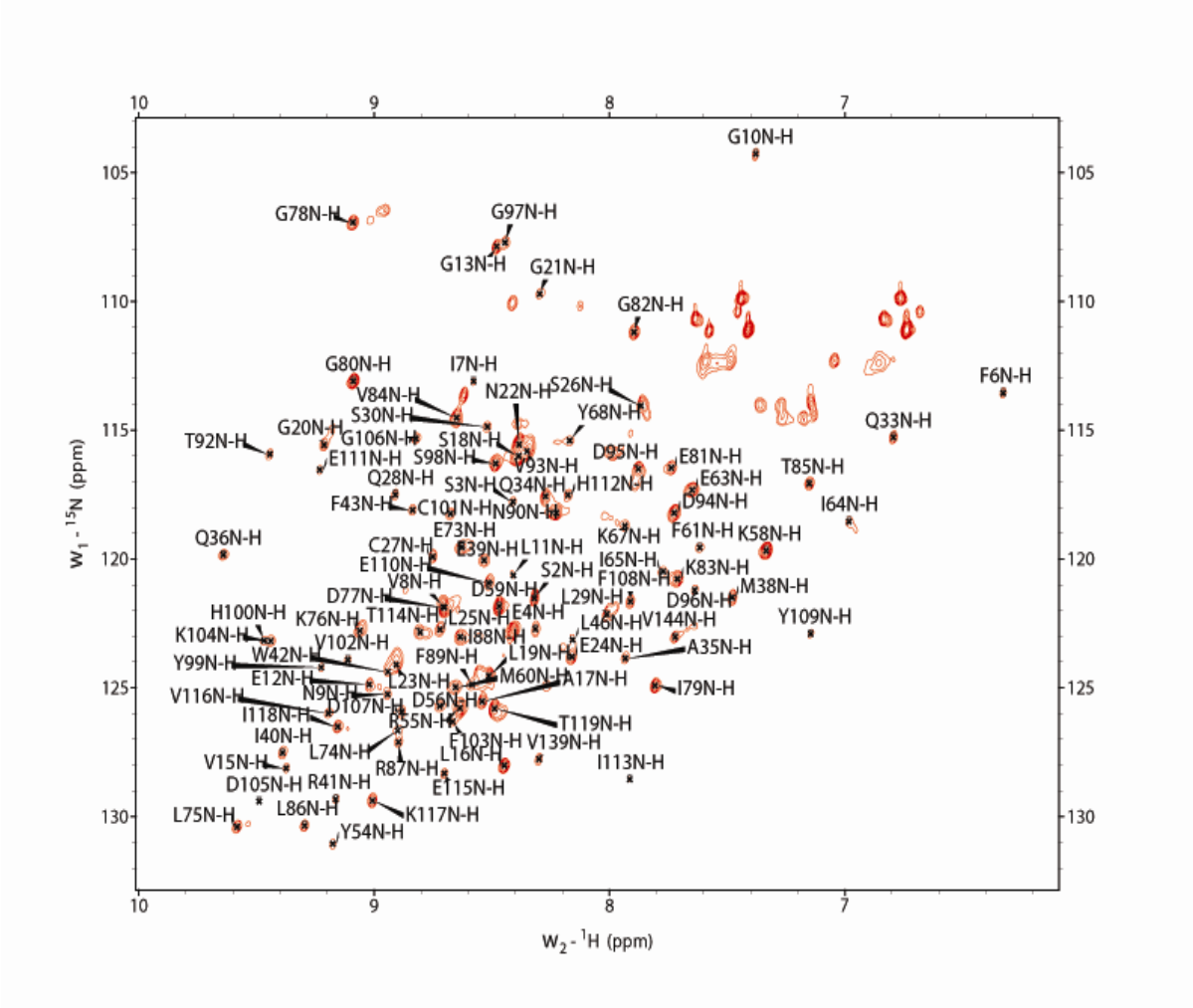


Figure 5.18 ^{1}H ^{15}N-HSQC spectrum of Skint-1 IgV. A standard ^{1}H ^{15}N-HSQC displaying all assigned backbone amide residues for Skint 1 IgV. Spectrum recorded on a varian Inova 800 MHz spectrometer and processed using NMRPipe (Delaglio F 1995) and analysed using SPARKY (Goddard TD 2004).

Residue Number	Residue type	Chemical Shift (ppm)				
		HN	N	Co	Cα	Cβ
1	Met	-	-	177.2	-	-
2	Ser	8.322	121.6	174.9	58.57	63.72
3	Ser	8.412	117.9	174	58.37	63.73
4	Glu	8.319	122.9	174.5	54.36	29.77
5	Pro	-	-	175.9	64.04	32.27
6	Phe	6.322	113.7	174.1	55.35	41.34
7	Iso	8.579	113.3	174.4	60.17	41.42
8	Val	8.711	121.9	174.5	61.63	33.52
9	Asn	8.944	125.4	174.2	50.92	41.92
10	Gly	7.379	104.4	172.3	46.17	-
11	Leu	8.418	120.6	178.8	55.12	42.15
12	Glu	9.011	124.9	178.2	58.02	30.18
13	Gly	8.479	108	178.1	44.27	-
14	Pro	-	-	176.2	61.8	34.09
15	Val	9.38	128.2	174.3	62.65	33.06
16	Leu	8.451	128.2	177	53.91	42.11
17	Ala	8.542	125.6	175.3	50.58	22.71
18	Ser	8.387	116.1	174.2	57.17	64.89
19	Leu	8.518	124.7	178.5	57.01	42.16
20	Gly	9.215	115.7	174.7	45.41	-
21	Gly	8.304	109.8	171.4	44.02	-
22	Asn	8.39	115.6	174.8	51.32	40.07
23	Leu	8.91	124.2	174.2	53.56	46.6
24	Glu	8.16	123.9	174.8	54.68	32.01
25	Leu	8.724	122.9	174.9	53.1	46.9
26	Ser	7.881	114.3	173.7	57.26	65.94
27	Cys	8.751	120	172.4	55.15	50.61
28	Gln	8.917	117.6	174	53.65	31.93
29	Leu	8.012	122.2	177.8	54.63	44.54
30	Ser	8.525	115	173.2	54.86	64.72
31	Pro	-	-	-	64.11	31.71
32	Pro	-	-	176.2	64.32	32.2
33	Gln	6.8	115.4	172.2	54.16	31.79
34	Gln	8.276	117.7	177.2	57.42	30.29
35	Ala	7.934	123.9	178.4	51.04	21.1
36	Gln	9.65	120	177.6	59.44	27.83
37	His	-	-	175.5	56.58	30.2
38	Met	7.48	121.6	174.4	57.4	32.61
39	Glu	8.535	120.2	175.3	54.91	31.96
40	Iso	9.391	127.7	174.8	60.13	40.69

41	Arg	9.164	129.4	173.8	54.26	36.04
42	Trp	8.94	124.4	176.9	54.57	31.81
43	Phe	8.839	118.2	171.8	56.63	41.32
53	Leu	-	-	171.3	54.98	45.97
54	Tyr	9.179	131.2	173.5	55.81	42.26
55	Arg	8.637	125.9	174.7	56.09	33.58
56	Asp	8.73	125.8	175.9	55.26	40.03
57	Gly	-	-	174.1	45.29	-
58	Lys	7.337	119.8	174.4	54.57	35.94
59	Asp	8.469	121.9	176.1	54.85	42.13
60	Met	8.658	125.1	176.7	52.36	29.62
61	Phe	7.621	119.7	177.8	61	39.29
62	Gly	8.961	106.5	174.7	46.25	-
63	Glu	7.647	117.4	175.6	55.8	31.09
64	Iso	6.984	118.7	176.2	61.2	38.75
65	Iso	7.777	120.6	178.2	62.39	38.35
66	Ser	-	-	176.5	62.16	62.95
67	Lys	7.938	118.8	176.3	58.25	33.16
68	Tyr	8.174	115.5	176.3	60.04	40.77
70	Glu	-	-	176.9	57.88	-
71	Arg	7.991	116	174.3	56.34	32.98
72	Thr	7.85	108.5	174.6	59.5	72.56
73	Glu	8.631	119.7	172.6	55.55	34.32
74	Leu	8.905	126.7	175.7	52.62	43.72
75	Leu	9.586	130.6	179.7	55.6	40.64
76	Lys	9.067	122.9	177.7	56.85	33.26
77	Asp	8.705	122.1	177.1	57.53	40.22
78	Gly	9.095	107.1	177.2	45.51	-
79	Iso	7.81	125	177.5	66.3	39.62
80	Gly	9.09	113.2	174.1	46.05	-
81	Glu	7.74	116.6	177.4	55.2	30.98
82	Gly	7.897	111.3	172.1	47.11	-
83	Lys	7.713	120.9	176.7	55.07	35.52
84	Val	8.656	114.6	173.8	59.93	36.27
85	Thr	7.156	117.2	171.6	61.25	72.1
86	Leu	9.298	130.5	173.5	53.17	43.2
87	Arg	8.902	127.3	174.3	54.03	33.23
88	Iso	8.637	123.2	175.4	59.79	40.45
89	Phe	8.586	125	176.1	56.82	41.1
90	Asn	8.23	118.3	174.7	53.57	36.73
91	Val	8.647	114.6	177.3	62.95	32.98
92	Thr	9.448	116.1	175.6	59.75	72.72
93	Val	8.355	116	177	65.09	31.61

94	Asp	7.724	118.3	176.1	55.35	41.08
95	Asp	7.88	116.6	176.2	55.01	41.51
96	Asp	7.641	121.3	175.7	55.04	42.31
97	Gly	8.448	107.8	172.8	45.17	-
98	Ser	8.484	116.4	174.4	58.11	63.23
99	Tyr	9.226	124.3	175.9	58.28	41.05
100	His	9.476	123.3	173.6	53.43	29.12
101	Cys	8.681	118.3	172	53.24	40.92
102	Val	9.113	124	175	60.73	35.13
103	Phe	8.662	126.4	173.9	56.52	41.27
104	Lys	9.44	123.3	175	54.85	37.5
105	Asp	9.49	129.5	176.2	52.47	43
106	Gly	8.825	115.4	174.6	47.23	-
107	Asp	8.886	126.1	175.7	54.96	41.32
108	Phe	7.914	121.8	173.3	58.24	41.43
109	Tyr	7.147	123	172.6	55.85	41.44
110	Glu	8.52	121.1	174.4	55.54	36.5
111	Glu	9.241	116.7	174.7	54.31	36.29
112	His	8.18	117.6	172.1	55.28	31.55
113	Iso	7.912	128.6	174.3	60.79	39.78
114	Thr	8.804	122.9	173.1	61.99	71.75
115	Glu	8.703	128.4	175.1	56.47	30.98
116	Val	9.196	126.1	175.9	61.3	35.2
117	Lys	9.009	129.5	174	54.44	35.56
118	Iso	9.167	126.6	176.5	58.98	37.62
119	Thr	8.491	125.9	178.7	62.76	71.37
121	Val	-	-	175.8	63.44	32.07

Table 5.2 Chemical shift assignments for Skint-1 IgV. *Chemical shifts were recorded from the HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH, HNCO and HN(CA)CO experiments with $^{15}N/^{13}C$ labelled Skint-1 IgV protein.*

5.5.4 Assignment of Skint-1 IgV side chains

The majority of the side chain resonances of the Skint-1 IgV protein were identified by the analysis of the HCCH-TOCSY, (H)CCH-TOCSY, HCCONH, CCONH and ^{13}C-edited aromatic NOESY spectrum. The HCCONH and CCONH experiments provided a lower resolution map of resonance assignments in the side chain correlated to the backbone amide

(Figure 5.19). This provided a key starting point for the assignment of the more complex but high resolution TOCSY experiments which allowed the assignment of the aliphatic side chains (Figure 5.20). The NOESY spectrum provided through space interactions within 5Å and was predominantly used for confirming assignment locations and aromatic side chains allowing correlation back to the side chain aliphatic residues in close proximity to the aromatic moiety.

In total 84.9% of H^{α}, H^{β}, H^{γ}, H^{δ} and H^{ε}, 89.9% of C^{α}, 80.7% of C^{β}, 73.5% of C^{γ}, 85.7% of C^{δ}, and 100% of C^{ε} aliphatic resonances were assigned. The ^{1}H and ^{15}N resonances for the NH_2 side chains for 3 out of 4 Asparagine and 4 Glutamine were assigned completely. No assignments were made for the labile guanidine moiety of Arginine, the side chain NH_3^+ of Lysine, and the side chain carboxyl groups of Aspartate and Glutamate (see Appendix II).

5.5.5 Secondary structure predictions

The secondary structure of the Skint-1 IgV domain was predicted using the chemical shift difference method between measured values and random-coil values of C^{α}, C^{β} and (ΔC^{α} - ΔC^{β}) based on the TALOS (Cornilescu *et al*, 1999) and CSI (Wishart and Sykes, 1994) protocols. The comparisons between folded and random coil chemical shifts show that Cα resonances tend to shift upfield in β-sheets and extended stands, but tend to shift downfield in α-helices. The reverse is true for the Cβ resonances. As the Cα and Cβ secondary shifts are of similar magnitude and opposite sign for both helices and shifts, subtraction of the Cα and Cβ secondary shifts (ΔC^{α} - ΔC^{β}) enhances the correlation between the secondary structural elements and the secondary shifts. The ΔC^{α} - ΔC^{β} and TALOS plots represent an empirical prediction of a protein's phi and psi backbone torsion angles using HN, HA, CA, CB, CO and

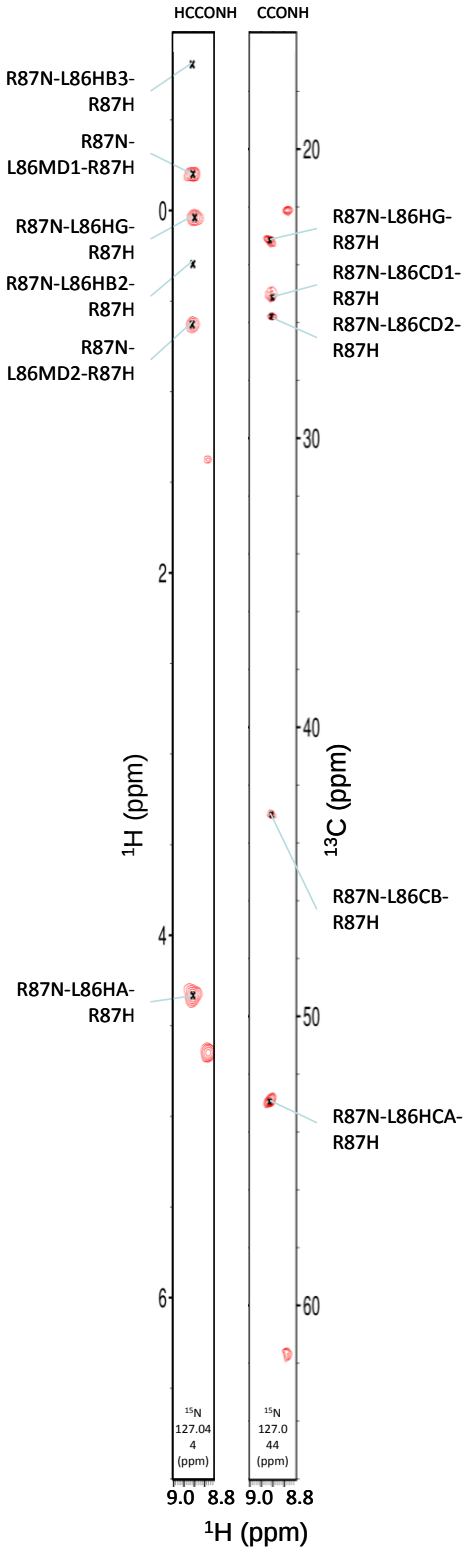


Figure 5.19 HCCONH and CCONH strip plots at a ^{15}N chemical shift of 127.0 ppm showing the side chain assignments for residue Leucine 86. The Spectra were recorded at 303 K on a Varian Inova 800 MHz NMR spectrometer equipped with triple resonance cryogenic probe and z-axis pulse field gradients

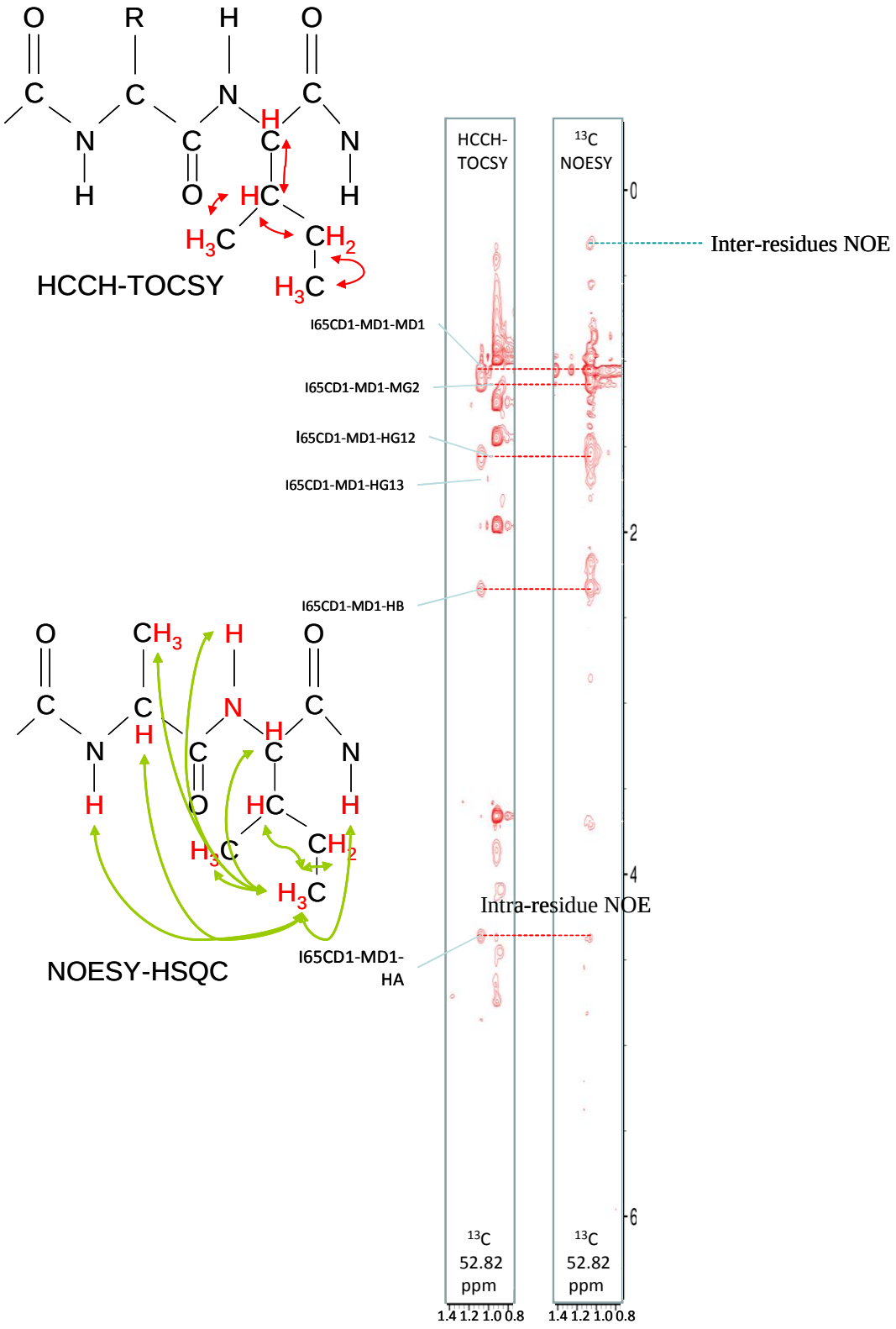


Figure 5.20 CD1-MD1 TOCSY and NOESY interactions for I65 displayed in two strip plots. Protons up to the Hγ can be identified and intra residue NOE's and inter residue NOE's are visible and allow for the sequential assignment and structure determination of the protein. The red arrows in the HCCH-TOCSY show the through bond direction of magnetism. The green arrows in the NOESY-HSQC show the through space direction of magnetism displaying the possible inter-residue assignments to within a distance of 5Å.

N chemical shift assignments for a given residue (Figure 5.21). The CSI program for determining secondary structure in proteins from the chemical shift indices of ^{1}H and ^{13}C nuclei plot indicated the presence of 3 α-helices and 8 β-strands of at least 3 amino acids in length (Figure 5.21); the α-helices were estimated to encompass residues I64 to Y68, D77 to E81 and T92 to G97 whilst the β-strands span residues V7 to G9, V15 to A17, G21 to Q28, E39 to F43 and L53 to R55, T72 to L75, K83 to F89, S98 to D105 and F108 to I118. The assignments and secondary structure prediction of the Skint-1 IgV domain provided the basis for the characterisation of its three dimensional structure and interaction with other cellular molecules.

5.5.6 Hydrogen deuterium (H-D) exchange

The Hydrogen deuterium (H-D) exchange method provides data regarding hydrogen bonding and the solvent accessibility of different parts of the protein, and thus provides key restraints for structure determination. The principle is based on measuring the rate of exchange of a protonated protein with that of a deuterated solvent. For example, an amide which contributes to an intramolecular hydrogen bond will exchange very slowly, whereas an amide located on the surface of protein will exchange more rapidly. In addition, amides buried from the solvent but not mediating hydrogen bonding interactions may also exhibit slower exchange rates. Therefore, when a protein is transferred from a protonated to a deuterated solvent, only signals from residues protected from the solvent, remain.

Since the nuclei of hydrogen and deuterium exhibit significantly different magnetic properties it is possible to discriminate between them using NMR spectroscopy. Generally, HSQC spectra are measured at a number of time intervals during which the H-D exchange is taking place. As the HSQC experiment is specific for hydrogen, the signal decays exponentially as

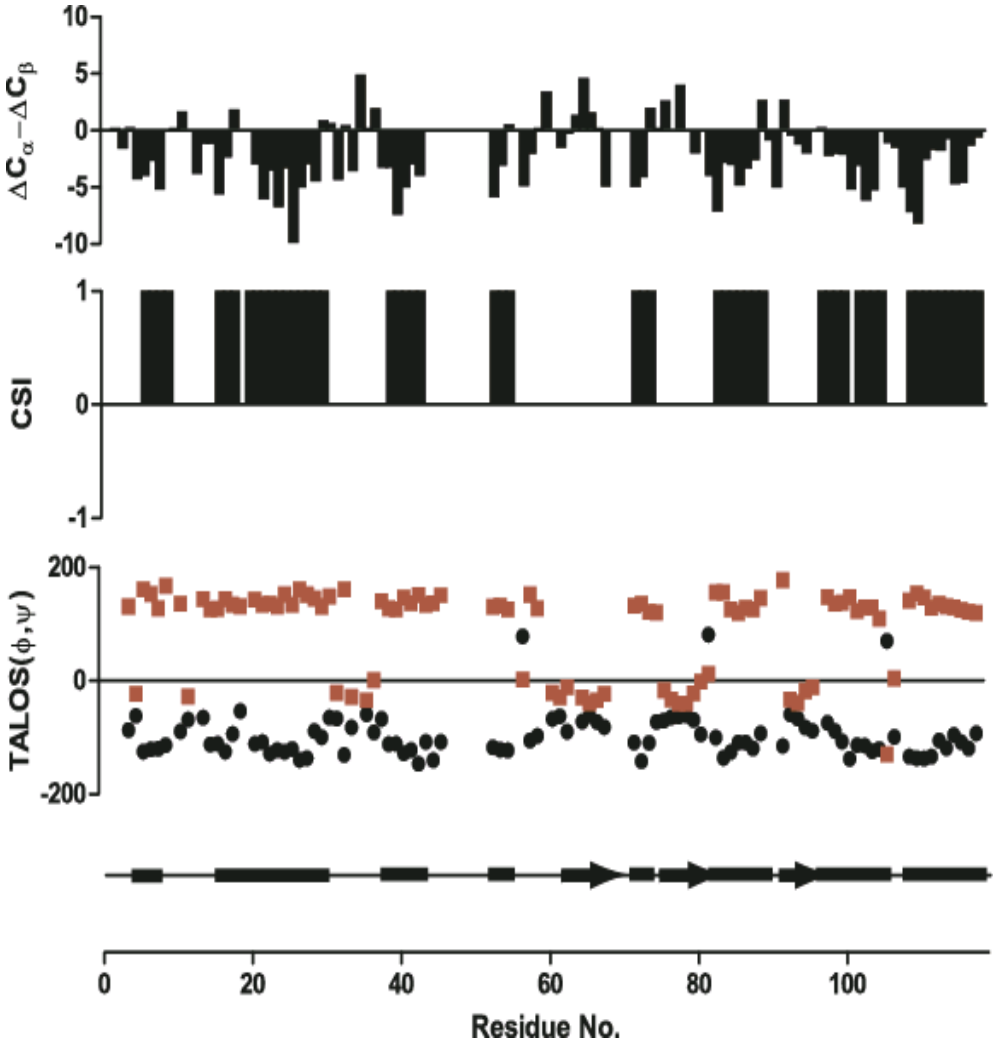


Figure 5.21 Summary of secondary structural predictions of Skint 1 IgV. (ΔCα - ΔCβ) highlights the enhanced correlation between the secondary structural elements and the secondary shifts, CSI and TALOS results plotted. In the consensus CSI, the values '1' represents the β-strand tendency, while '-1' represents α -helical tendency, the lack of α helices maybe due to their size being only 3^{10} helices and undetectable by CSI. Backbone dihedral angles (phi, psi) were calculated using TALOS. Phi (/) and psi (w) angles are shown in black and red, respectively. Predicted secondary structure elements are indicated using rectangles for β-strands and arrows for α-helices

the hydrogen exchanges. This experiment provides amino acid-specific data for all the residues in the protein simultaneously. To establish which Skint-1 IgV residues were buried or participating in hydrogen bonding interactions, HSQC experiments were carried out with samples before and one hour after incubation with deuterium oxide (Figure 5.22). The resulting HSQC spectrum demonstrated that ~63% of the residues in the protein exchanged hydrogen for the deuterated solvent. This indicates that the remaining 37% (42 residues) are protected from exchange and hence are either completely buried or most likely involved in hydrogen bonding (Table 5.3).

Residue number	**Amino acid**		**Residue number**	**Amino acid**
17	Ala		85	Thr
23	Leu		86	Leu
25	Leu		87	Arg
27	Cys		88	Ile
29	Leu		89	Phe
38	Met		90	Asn
39	Glu		96	Asp
40	Ile		99	Tyr
41	Arg		100	His
43	Phe		101	Cys
54	Tyr		102	Val
55	Arg		103	Phe
73	Glu		104	Lys
75	Leu		105	Asp
79	Ile		110	Glu
81	Glu		112	His
82	Gly		114	Thr
83	Lys		116	Val
84	Val		117	Lys

Table 5.3 Hydrogen-deuterium (H-D) exchanged residues. *Highlighted are the amides involved in hydrogen deuterium exchange. These residues were protected from the hydrogen deuterium exchange process therefore indicating that they are either involved in hydrogen bonding or buried in the core of the protein.*

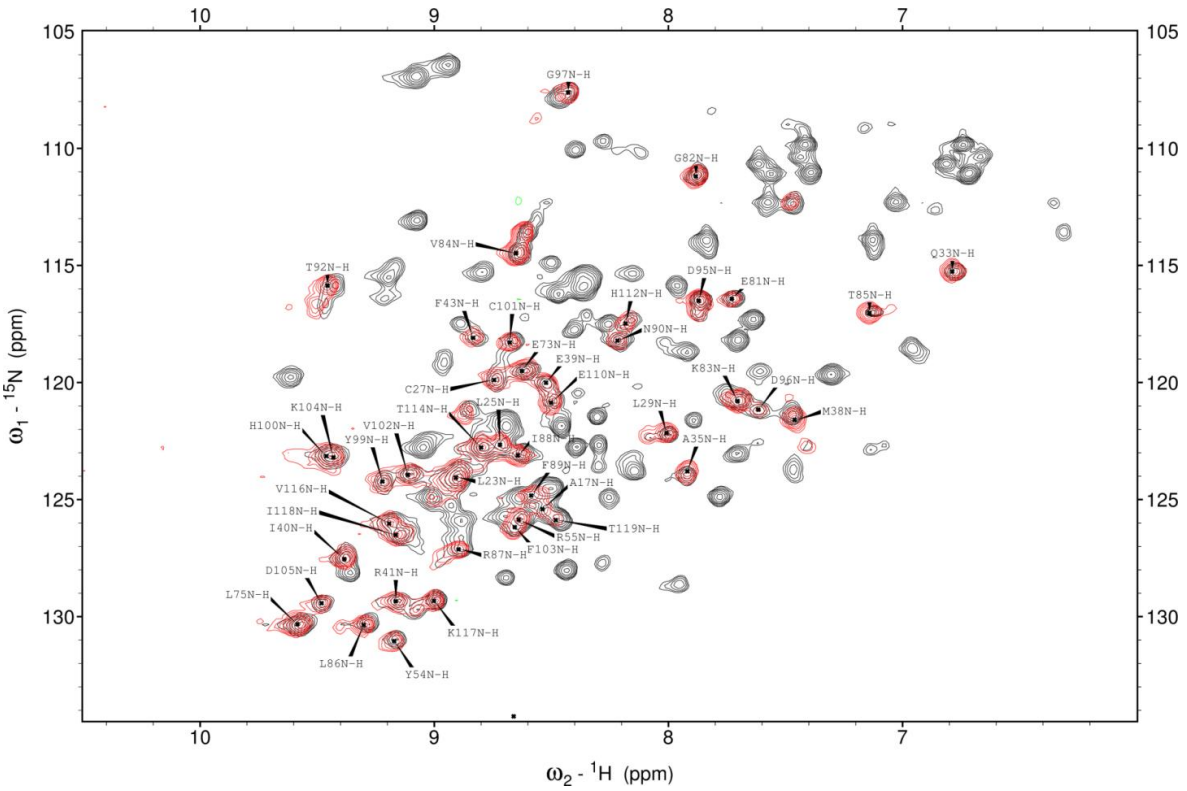


Figure 5.22 Skint-1 IgV H/D exchange analysis. ^{15}N Skint-1 IgV in 20mM MES pH 6.5, 50mM NaCl in 100% H_2O buffer (Black) and following 1 hour incubation in 100% D_2O buffer (Red). Those resonances remaining after 1 hour are labelled.

5.5.7 Structure calculation

The 3D structure of Skint-1 IgV domain was determined using the CANDID/CYANA package with automated NOE cross peak assignment and structure calculation with torsion-angle dynamics implemented (Güntert, 2004). Three NOE spectra were used to provide restraints, the ^{15}N-edited NOESY-HSQC, the ^{13}C-NOESY-HSQC and the ^{13}C-aromatic-NOESY-HSQC, together with dihedral angles calculated using TALOS. The protocol consisted of seven iterative cycles of NOE assignment and structure calculation. The first cycle started from 100 random conformers with 10000 torsion angle dynamic steps performed per conformer in each cycle and the results inputted into the next cycle. 20 conformers with the lowest CYANA target function were selected as representative. Using the initial structure as a template together with the H/D exchange data, hydrogen bonds were mapped and the structure was further refined.

A total of 894 NOEs, 38 hydrogen bonds and 182 dihedral angle restraints were used in the final calculation. Structures were analysed using PROCHECK-NMR. The RMSD of the 20 lowest energy structures was calculated using MOLMOL. The quality parameters of the Skint-1 IgV structure ensemble are listed in Table 5.4. Despite a significant number of peaks lacking from the ^{1}H-^{15}N-HSQC spectrum, almost complete resonance assignment of the side chain signals was attained. Hence the structure determination of Skint-1 IgV was successful and produced a high quality structure. All distance restraint violations fall under 0.2Å. On average also the dihedral angle violations are small: in 20 substructures only two violations exceed 10°. Both backbone and heavy atom RMSD values of the structure family are very low and coordinate precision is uniform throughout the structure. There are 1.3% residues in

the disallowed regions of the Ramachandran plot. The 20 structures for the Skint-1 IgV molecule were visualised utilizing PyMol (Figure 5.23)

Number of restraints	
NOE	
Total	894
Intraresidue (i=j)	117
Sequential (|i-j|)=1	194
Medium range (1 < |i-j|≤4)	140
Long range (|i-j| > 4)	443
Hydrogen bond restraints	38
Dihedral-angle restraints	182
RMSD from distance restraints (Å)	0.018 ± 0.002
RMSD from covalent geometry	
Bond lengths (Å)	0.004 ± 0.000
Angles (°)	0.509 ± 0.019
Impropers (°)	1.517 ± 0.123
Ramachandran plot (%)	
Most favoured regions	80.5
Additionally allowed regions	16.5
Generously allowed regions	1.6
Disallowed regions	1.3
RMSD (Å)	
Backbone (residues 8-118)	1.42
Backbone (Absent undefined regions 8-14 & 44-71)	0.83
Heavy atoms (residues 8-118)	2.28
Heavy atoms (Absent undefined regions 8-14 & 44-71)	1.58

Table 5.4 Structural statistics for Skint-1 IgV (20 structures). *The statistics for the Skint-1 IgV ensemble were calculated using PROCHECK-NMR.*

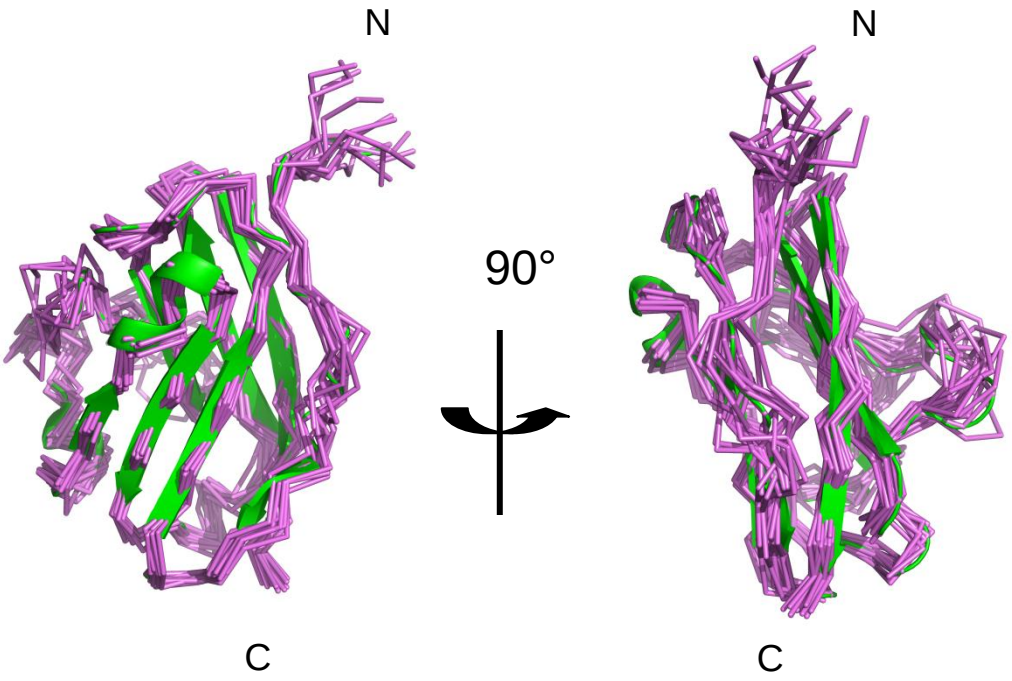


Figure 5.23 Solution structure of mouse Skint-1 IgV domain highlighting the 20 lowest energy states derived from NMR spectroscopic analysis. The lowest energy state is depicted in green (cartoon form).

5.6 Discussion

The structure determination of the Skint-1 ectodomain could provide a greater molecular understanding of its role in selection of the DETC subset, and also potentially antigen recognition by DETC resident in the skin. Unfortunately, the prokaryotic derived Skint-1 ectodomain inclusion body material proved intransigent to refolding, despite extensive attempts with conventional oxidative *in vitro* refolding strategies. This is particularly surprising given that there are a considerable number of studies highlighting the use of similar strategies to refold proteins encompassing two tandem Ig domains in sufficient levels for structural studies (Willcox *et al*, 2002; Nicholls *et al*, 2009 and Simpson *et al* 2011). There are several possible explanations for why the Skint-1 IgV-IgC construct proved resistant to refolding. Firstly, the constituents of the refolding mixture are most likely sub-optimal for efficient renaturation of the Skint-1 IgV-IgC domains. It is possible that the refolding efficiency of the Skint-1 ectodomain could be enhanced by varying constituents of the refolding buffer. This could involve pulsing with smaller quantities of inclusion body protein, sampling different denaturant buffers and varying the ratio of oxidised and reduced glutathione. Secondly, SDS-PAGE analysis of the Skint-1 IgV-IgC inclusion bodies demonstrated a high level of background proteins which may have interfered with the refolding process. Indeed, previous experiments have also demonstrated that impurities within the inclusion body material have detrimental effects on protein renaturation and to successfully overcome intransigent refolding additional purification of the inclusion bodies were required (Ouellette *et al,* 2003). A similar strategy could be employed for the Skint-1 IgV-IgC inclusion bodies to improve purity and enhance the efficiency of the renaturation process. However, one of the major problems associated with this approach is that the

recovery yield of the inclusion body material following additional purification steps is often extremely poor (~ < 10%) for subsequent refolding.

Finally, the domain boundaries selected for the Skint-1 ectodomain construct may have inadvertently affected the renaturation step. As mentioned previously, the secondary structure predictions of the Skint-1 molecule revealed three distinct domains; a membrane distal IgV domain followed by an IgC-like domain which then leads to a transmembrane spanning domain. The Skint-1 ectodomain construct used for structural studies encompassed the two tandem Ig domains but lacked the transmembrane spanning region. It was envisaged that by removing the transmembrane sequence it may promote Skint-1 crystallisation. However, it is possible that the transmembrane region is important for stabilising the IgC-like domain, and in its absence it led to inefficient refolding or yielded over-expressed protein that was very unstable. Indeed, previous secondary structure predictions have revealed that elements of the transmembrane region protrude outwards on the extracellular side. It is possible that some of these protruded regions may play a key role in stabilising the IgC domain. Hence, future Skint-1 constructs should factor into account the inclusion of the transmembrane spanning region. Alternatively, the IgV-IgC construct may be too long such that the residues at the C-terminal end prevent efficient refolding of the Skint-1 ectodomain. Future cloning strategies could involve utilising truncated versions of the IgV-IgC construct, whereby residues at the C-terminal end are systematically removed.

Since the Skint-1 IgV-IgC construct was intransigent to oxidative refolding, an alternative strategy involved shortening the construct in order to refold each domain independently. It was envisaged that of the two domains the Skint-1 IgV domain was most likely to play a

functionally significant role in selecting the DETC cells and hence was chosen for further characterisation. As demonstrated with LILRB4 (PhD Thesis, L.I. Garner, 2011), the single Skint-1 IgV domain construct was successfully refolded in sufficient quantities for structural characterisation.

Initial attempts at determining the three dimensional structure of Skint-1 IgV domain involved using X-ray crystallography. Following extensive crystallisation screening and optimisation process, suitably sized Skint-1 IgV crystals were grown. These crystals were reproducibly generated allowing multiple attempts to determine the optimum cryoprotectant conditions prior to X-ray data collection. Although the final optimised Skint-1 IgV crystals were of reasonable size, they diffracted X-rays to 4Å resolution with the 'in house' rotating anode generator. The data processing statistics calculated for the Skint-1 IgV crystal were deemed acceptable and did not highlight major causes for concern. Nevertheless, attempts at solving the Skint-1 IgV structure with molecular replacement using MOG as the search model proved to be unsuccessful despite the relatively high sequence homology between the two proteins. It was possible, but unlikely, that higher resolution data obtained from a synchrotron radiation source may enhance the likelihood of determining the crystal structure of Skint-1 IgV. Since molecular replacement failed to provide meaningful phase information it was theoretically possible to experimentally calculate phases using MAD or MIR. However, as both approaches were time-consuming and given that the native crystals diffracted to a low resolution, a new strategy was employed in which the Skint-1 IgV was subjected to NMR analysis. Indeed, previous studies have demonstrated that proteins encompassing Ig domain architecture were amenable to structural characterisation by NMR spectroscopy (Gallagher *et al*, 1994).

Preliminary assessment of the Skint-1 IgV domain suggested that the protein exhibited the necessary properties for structure determination using NMR spectroscopy. In particular, the highly concentrated Skint-1 protein was relatively stable for long periods of time at room temperature. Initial analysis revealed that although 80% of the backbone assignments could be assigned the remaining could not due to lack of chemical shift resonances. This had significant ramifications in terms of resolving the overall Skint-1 structure as the backbone experiments were expected to provide information regarding the overall basic fold of the molecule. To supplement these data, further experiments involving ^{15}N-edited HCCH-TOCSY, (H)CCH-TOCSY, HCCONH, CCONH were performed to allow side chain assignments to be resolved, thereby decreasing the level of unassigned residues to approximately 15%. Despite these efforts, there still remained a significant proportion of residues that were unassigned most likely due to these residues conforming to a less stable and dynamic structure. In future to overcome the lack of resonances in this region the residues could be mutated, thus stabilising this region, therefore allowing chemical shift data to be recorded. However given the large proportion of unassigned residues, this approach could be problematic as it may affect functionally important regions of the Skint-1 IgV domain. Alternatively, stabilising compounds such as trimethylamine *n*-oxide (TMAO), a naturally occurring osmolyte that stabilizes proteins, could be utilised to enhance the rigidity and facilitate the stabilisation of the dynamic region. It is possible that these highly flexible regions prevented tight packing within the Skint-1 IgV crystal lattice, which contributed to the Skint-1 IgV crystals providing relatively low resolution X-ray data.

Since the structure of the Skint-1 IgV domain has successfully been determined with NMR, it opens up the possibility of using a similar strategy to resolve the structure of the IgC-like

domain. This would then allow for a model of the IgV-IgC structure to be constructed and may provide a more comprehensive understanding of Skint-1 ligand binding sites and/or partners. The analysis of the Skint-1 IgV structure is presented in significant detail in Chapter 6.

CHAPTER 6 –

SOLUTION STRUCTURE ANALYSIS OF THE SKINT-1 IgV DOMAIN

Chapter 6: Solution structure analysis of the Skint-1 IgV domain

6.1 Overall Structure of Skint-1 IgV

NMR derived structure calculations produced 20 solution models for the structure of the mouse Skint-1 Immunoglobulin variable (IgV) domain (Figure 6.1). The final structural ensemble is well-defined, yielding a mean backbone root-mean-square deviation (RMSD) of 0.83 Å for the structured regions. Structural analysis reveals the Skint-1 IgV domain adopts an immunoglobulin (Ig) fold, consisting of nine β stands, with three 3^{10} α helices that are positioned between the C'-D, D-E and E-F strands (see Table 6.1).

Residue number	Secondary structure	Residue number	Secondary structure
S1 to I6	Random coil	V69 - R71	Random coil
V7 to G9	β Strand A	T72 to L75	β Strand D
L10 to P14	Random coil	K76	Random coil
V15 to A17	β Strand A'	D77 to E81	3^{10} helix 2
S18 to G20	Random coil	G82	Random coil
G21 to Q28	β Strand B	K83 to F89	β Strand E
L29 to M38	Random coil	N90 - V91	Random coil
E39 to R44	β Strand C	T92 to G97	3^{10} helix 3
N45 to H52	Random coil	S98 to D105	β Strand F
L53 to R55	β Strand C'	G106 - D107	Random coil
D56 to F63	Random coil	F108 to I118	β Strand G
I64 to Y68	3^{10} helix 1	T119	Random coil

Table 6.1. Secondary structure assignments for the Skint-1 IgV domain derived from the solution structure

Consistent with its homology to Ig variable-type domains, the Skint-1 IgV structure adopts a compact beta sandwich domain comprising two anti-parallel sheets (one consisting of strands A, A', B, D, E, the second C, C', F, G), with the N and C termini at opposite ends of the molecule (Figure 6.2). A large hydrophobic core, formed by in-pointing hydrophobic residues from the B (L23, L25 and L29) and E (V84 and L86) strands in sheet 1 and residues from the C (I40, W42 and F43) and F (Y99 and F103) strands in sheet 2, contributes to stabilising the interior of the IgV domain (Figure 6.3). The Skint-1 IgV domain is further stabilised by an

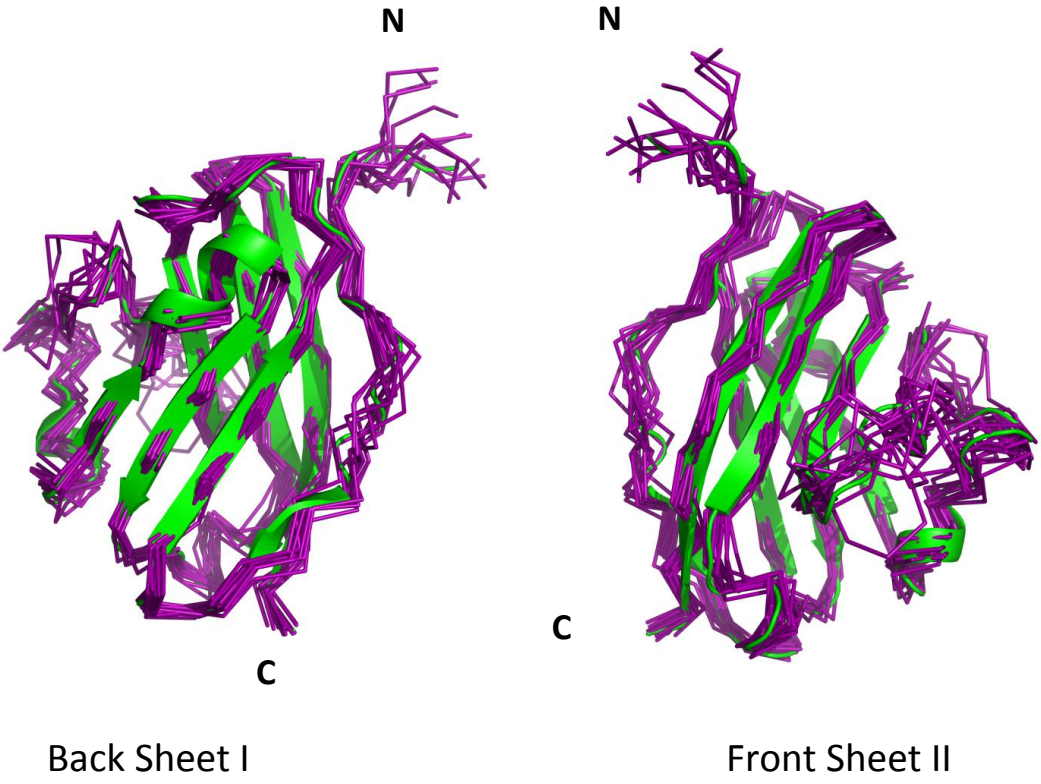


Figure 6.1 20 NMR structure solutions of the membrane-distal Skint-1 IgV domain with each state displayed in a simple ribbon (in purple). The lowest energy state is highlighted as a cartoon (in green) and was used in the subsequent structural analysis.

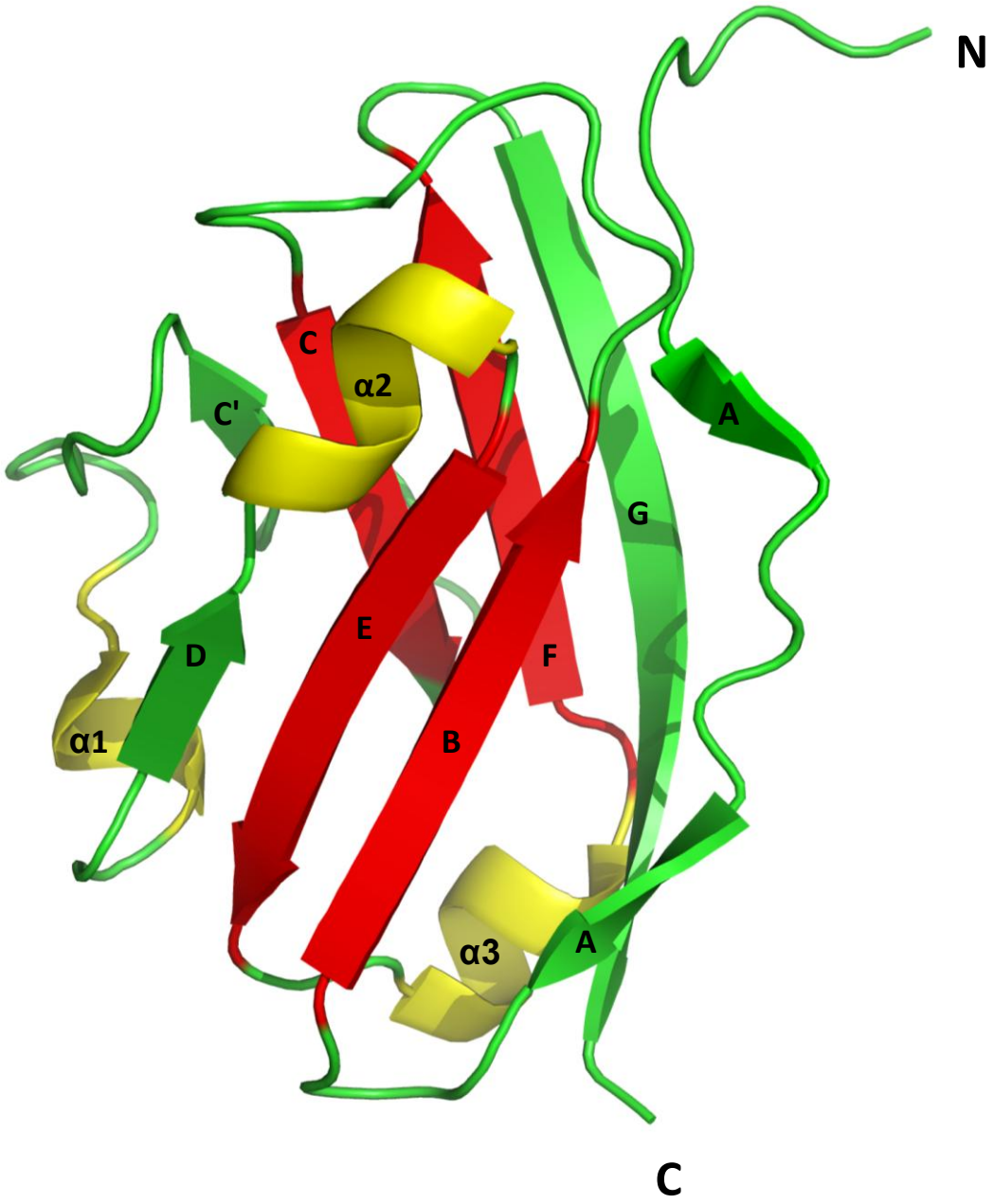


Figure 6.2 Ribbon representation of the lowest energy NMR structure of Skint1 IgV. The strands highlighted in red represent the core region of a typical IgV-type fold. Helices are shown in yellow. The β-strands are labelled according to the IgV-type fold described by Bork *et al*, 1994.

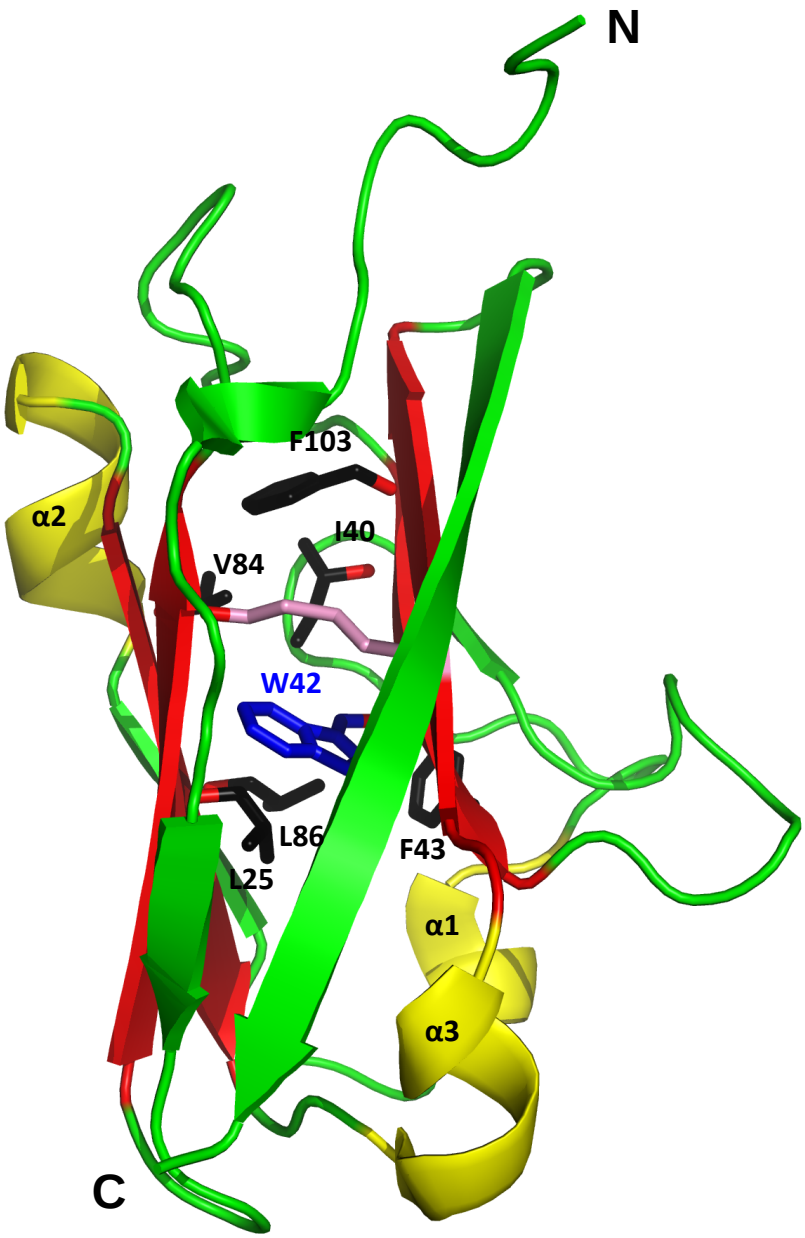


Figure 6.3 Stabilisation of the Skint-1 IgV core. The beta sheets are stabilised by a inter-chain disulphide bond (mediated by C27 & C101; shown in pink), an invariant tryptophan residue (W42; depicted in blue) and multiple non-polar interactions mediated by hydrophobic residues (shown as black sticks) from the core B, E, F and C strands (highlighted in red). Helices are shown in yellow.

intra-chain disulphide bond formed between strands B and F (mediated by residues C27 and C101) (Figure 6.3). This prominent feature, which connects the beta sheets, is similar to other IgV-like domain containing proteins and is a structural hallmark of Ig-folds (Lesk & Chothia, 1982). Also, there are 74 residues found between the conserved disulphide bond in Skint-1 IgV, which is reminiscent of previously reported variable and variable-related domain structures (65-75 residues) (Williams and Barclay, 1988). As reported with other Ig domains, the internal core region of the Skint-1 IgV domain is additionally stabilised by an invariant tryptophan residue (W42) packing against the disulphide bond (Figure 6.3).

The hydrogen-deuterium exchange analysis demonstrates that although the A to A' strand region is structurally defined via NOE restraints, it is weakly associated with the structure, as no protection of exchangeable amides was noted. In addition, the A, A', C and C' β-sheet regions show reduced conformational rigidity resulting from a lack of NOE restraints, most likely due to dynamics present within these regions, consistent with previous observations of IgV-domain structures (Kabsch and Sander, 1983).

6.2 Comparisons of Skint-1 with structural homologues

Automated structural comparisons using the DALI software revealed that the Skint-1 IgV structure is closely related to other IgV domains, in particular murine myelin oligodendrocyte (MOG) (PDB code 1PY9) and human Programmed Death-1 Ligand 1 (PD-L1) (PDB code 3FN3), with Z scores (corresponding to strength of structural similarity) of 13.8 and 12.7, respectively.

MOG is a type I transmembrane myelin protein of unknown function which is highly conserved between mouse and human, and is expressed in the Central Nervous System myelin. An autoantigen linked to pathogenesis of multiple sclerosis (MS) and experimental autoimmune encephalomyelitis (EAE), MOG is a minor component of the myelin sheath. In animal models, immunisation with MOG or MOG-specific T cells or even autoantibodies against MOG can lead to the development of neurological disease that mimics MS disease (Clements *et al*, 2003). Furthermore, MOG is commonly used to induce a demeylinating antibody response in mice with EAE as a model for MS (Breithaupt *et al*, 2008; Clements *et al*, 2003).

PD-L1 is a type I transmembrane protein that plays a major role in suppressing the immune system. The receptor for PDL-1 has been identified as Programmed Death 1 (PD-1) which is expressed on activated T cells and is a member of the CD28/CTLA-4 family. The formation of the PD-1/PD-L1 ligand complex is responsible for delivering an inhibitory signal which down-regulates T cell function, and maintains central and peripheral tolerance (Keir *et al*, 2008; Chen *et al*, 2010).

The primary amino acid sequences of murine Skint-1, murine MOG and human PD-L1 IgV domains were aligned with ClustalW2 (Figure 6.4). The alignments between murine Skint-1 and MOG demonstrated moderate levels of sequence conservation (sequence identity of 47% and overall sequence homology of 73%). In contrast, comparison of sequences between murine Skint-1 and human PD-L1 highlighted considerably less sequence conservation with a sequence identity of 24% (sequence homology of 45%). Notably, despite displaying modest to low levels of sequence conservation, both structural homologues adopt an IgV-type fold

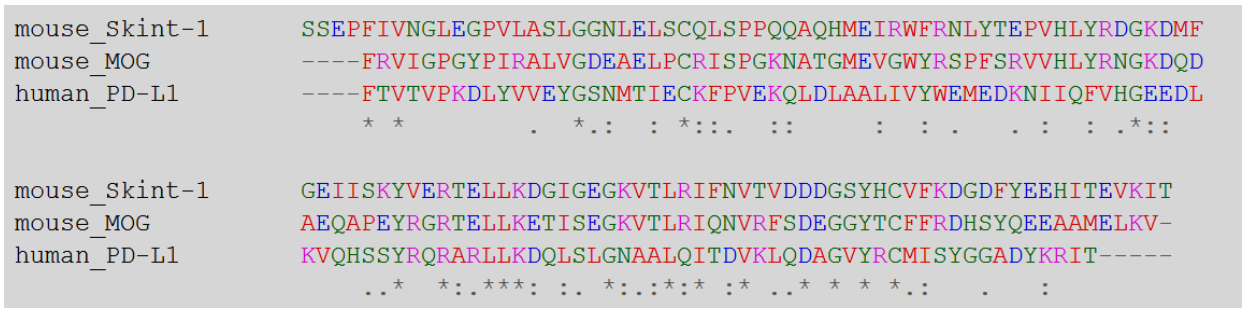


Figure 6.4 Multiple sequence alignment of the IgV domains of murine Skint-1, murine MOG and human PD-L1. Sequences of IgV domains were downloaded from the Uniprot database and aligned with ClustalW2. *Indicates residues that fully conserved, represents conservation between groups of strongly similar properties and indicates conservation between groups of weakly similar properties. Amino acids shaded in colour based on their chemical properties; red (small and hydrophobic), blue (acidic), magenta (basic) and green (hydroxyl, sulphydryl and amine).

(Clements *et al*, 2003 and Chen *et al*, 2010) and the structure of each can be superimposed onto Skint-1 IgV, resulting in an RMSD of 2Å (over 103 Cα aligned atoms) for MOG and 2.75Å for PD-L1 (over 92 Cα aligned atoms) (Figures 6.5-6.6). Furthermore, the majority of the internal facing non-polar residues that stabilize the Skint-1 Ig-V core are highly conserved in MOG and PD-L1.

6.3 Structural comparison of Skint-1 with MOG and PD-L1

Comparisons of the Skint-1 IgV domain with MOG demonstrates moderate levels of structural similarity as reflected by the RMSD value of 2Å (Figure 6.5). Unsurprisingly, regions that are most structurally compatible between MOG and Skint-1, particularly the strands that contribute to stabilising the IgV core, correlate with high levels of amino acid sequence identity (Figure 6.4). The majority of the strands, especially the A and B strands, are comparable in length between Skint-1 and MOG. Moreover, there are three additional segments that are highly conserved, both in terms of sequence and secondary structure, namely the C-C' loop (which adopts a random coil) and strands D and E (Figure 6.5).

Despite these similarities there are several topological differences observed between the Skint-1 and MOG IgV domains (Figure 6.5). With the exception of the loops connecting the strands, the most distinct structural features are restricted to two areas. Firstly, both the C and C' strands of MOG (E36-S42 and V48-Y51, respectively) are longer relative to Skint-1 (E39-R44 and L53-R55, respectively) (Figure 6.5). As a result, the interconnecting C-C' loop is considerably larger in Skint-1 relative to MOG and this may adversely affect the potential of Skint-1 to dimerise (more details below). Secondly, there are structural differences following the C' region, which in MOG includes a short β strand termed C'' (N55-D56), and two 3^{10}

Figure 6.5 Comparison of the lowest energy NMR structure of Skint-1 IgV with mouse myelin oligodendrocyte glycoprotein (MOG). The IgV domain of MOG (teal; PDB ID 1PY9, Clements *et al*, 2003) was superimposed onto Skint-1 (green) using align with a RMSD of 2Å (over 103 Cα aligned atoms). Secondary structural elements that differ between Skint-1 and MOG are labelled in red.

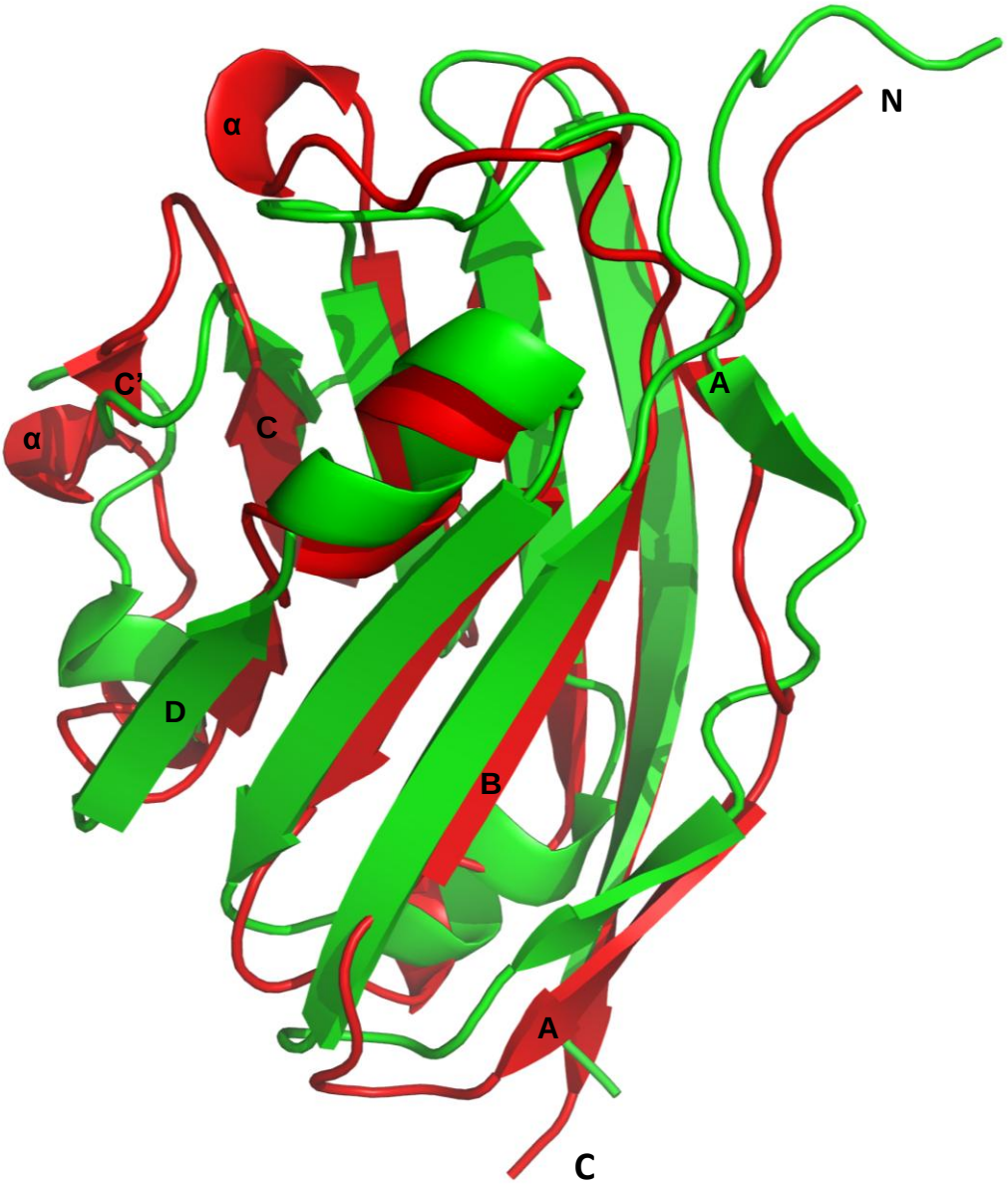


Figure 6.6 Comparison of the lowest energy NMR structure of Skint-1 IgV with human PD-L1. The IgV domain of PD-L1 (red; PDB ID 3FN3; Chen *et al*, 2010) was superimposed onto Skint-1 (green) using align with a RMSD of 2.75Å (over 92 Cα aligned atoms). Secondary structural elements that differ between Skint-1 and PD-L1 are labelled in black.

helical sections. In contrast, in Skint-1, the corresponding region lacks the C'' strand, and retains only one 3^{10} helix which may contribute to the increased flexibility observed relative to MOG. The lack of the C'' strand in Skint-1 is particularly surprising given the fact that, in this region, the primary amino acid sequence is relatively conserved with MOG (Figure 6.4).

Indeed, a paucity of NOE restraints for residues R44 to H52 of Skint-1 IgV provides further evidence for additional flexibility in this area. Moreover, temperature factor analysis of MOG highlights that the region following the C'' strand including the tandem 3^{10} helices are uniformly rigid. In addition, this region in MOG does not participate in any crystal contacts, suggesting that it is genuinely inflexible. One possible explanation for these differing observations may be attributed to the fact that the MOG structure represents a lower energy state relative to Skint-1, as MOG was resolved from a crystal at 100K by X-ray Crystallography, whereas the structure of Skint-1 IgV was determined in solution at 293K using NMR.

Despite the low amino acid sequence identity between PD-L1 and Skint-1 (24%), the structures of these domains possess remarkably similar conformations, particularly in the IgV core strands (Figure 6.6). Regardless of these structural similarities, there are several distinctive features that separate Skint-1 from PD-L1. Notably the PD-L1 IgV-like molecule contains 5 α helical regions, three of which are conserved in the Skint-1 domain; the additional α helical segments are located between the B-C (L50-A52) and C''-D (E72-A85) strands (Chen *et al*, 2010 and Figure 6.6). The equivalent regions in Skint-1 encompass random loop conformations and demonstrate considerable flexibility. Further topological differences include a shorter strand A in the PD-L1 structure relative to Skint-1. Moreover,

the A' (L27-E31) and C' (K62-F67) strands in PD-L1 are longer than the corresponding regions of Skint-1 (V15-A17 and L53-R55, respectively). However, strands B and D in PD-L1 (M36-K41, D85-L87) were found to be shorter relative to the equivalent regions of Skint-1 (G21-Q28, R71-L75) (Figure 6.6).

It is possible that these structural differences observed between Skint-1 and PD-L1 correlate with their unique biological function. However, this appears not to be the case based on detailed examination of the PD-1/PD-L1 complex structure. Indeed, the majority of PD-L1 regions that are critical for binding to and stabilising the PD-1 interface include the A, C, F and G strands (Lin *et al*, 2008). Strikingly, the majority of PD-L1 residues that mediate contacts to PD-1 are non-conserved in the corresponding regions of Skint-1, thereby providing a molecular basis for the non redundant roles of PD-L1 and Skint-1 despite adopting similar IgV folds.

6.3.1 Skint-1 shows radical alterations in the dimer interface region of MOG and PD-L1

Interestingly both structural homologues of Skint-1, MOG and PD-L1, exist as a dimeric species based on extensive biochemical and structural evidence (Clements *et al*, 2003 and Chen *et al*, 2010). Indeed, native MOG has previously been documented to exist as a mixture of monomeric and dimeric species (Clements *et al*, 2003). SDS-PAGE and Western blot analysis of myelin-derived proteins using a MOG-specific antibody resulted in detection of both monomeric and dimeric MOG protein (Clements *et al*, 2003). In addition, cross linking experiments with increasing concentrations of primary amine-coupling agent 3,3'-dithiobis[sulfosuccinimidylpropionate] (DTSSP), resulted in the detection of a MOG species

with a relative molecular mass equivalent to a dimer. Also, MOG adopted a head-to-tail dimer in the crystal lattice despite the size exclusion elution profile of renatured MOG demonstrating a predominantly monomeric species (Clements *et al*, 2003). This dimer interface is extensive, burying 1,800 $Å^2$ of surface area. Furthermore, the dimer interface exhibits a shape complementarity index of 0.7, consistent with a value that has been previously reported for antibody–antigen interactions (Clements *et al*, 2003). Taken together, this suggests that the crystallographic MOG dimer may represent a physiologically relevant MOG dimer (Clements *et al*, 2003).

Similarly to MOG, the PD-L1 structure also highlighted a dimer in the crystal lattice, with cross linking analysis performed on *in vitro* refolded protein confirming dimer formation (Chen *et al*, 2010). In contrast, the Skint-1 IgV domain behaves as a monomer in solution, as assessed by size exclusion chromatography, which highlighted an elution volume comparable to other known IgV-like monomeric proteins, such as β2-microglobulin (See Chapter 5 Figure 5.6). Furthermore, the solution structure of Skint-1 IgV determined by NMR revealed a stable monomeric protein.

Detailed structural analysis of MOG and PD-L1 has enabled identification of key residues that contribute to stabilising the dimer interface. For MOG, the dimer interface runs along the long axis of the protein, involving the A-A' loop, the C-C' loop and the F-G β hairpin (Figure 6.7a). More specifically, the MOG dimer interface is stabilised predominantly by van der Waal's and hydrogen bonding interactions mediated by G9, P11, S45, T97, Q106, E107, E108, A109, A110 and E112. Notably, whereas P11, E107, E108 and E112 of MOG are identical in Skint-1 IgV (P14, E110, E111, E115), the majority of the remaining amino acids

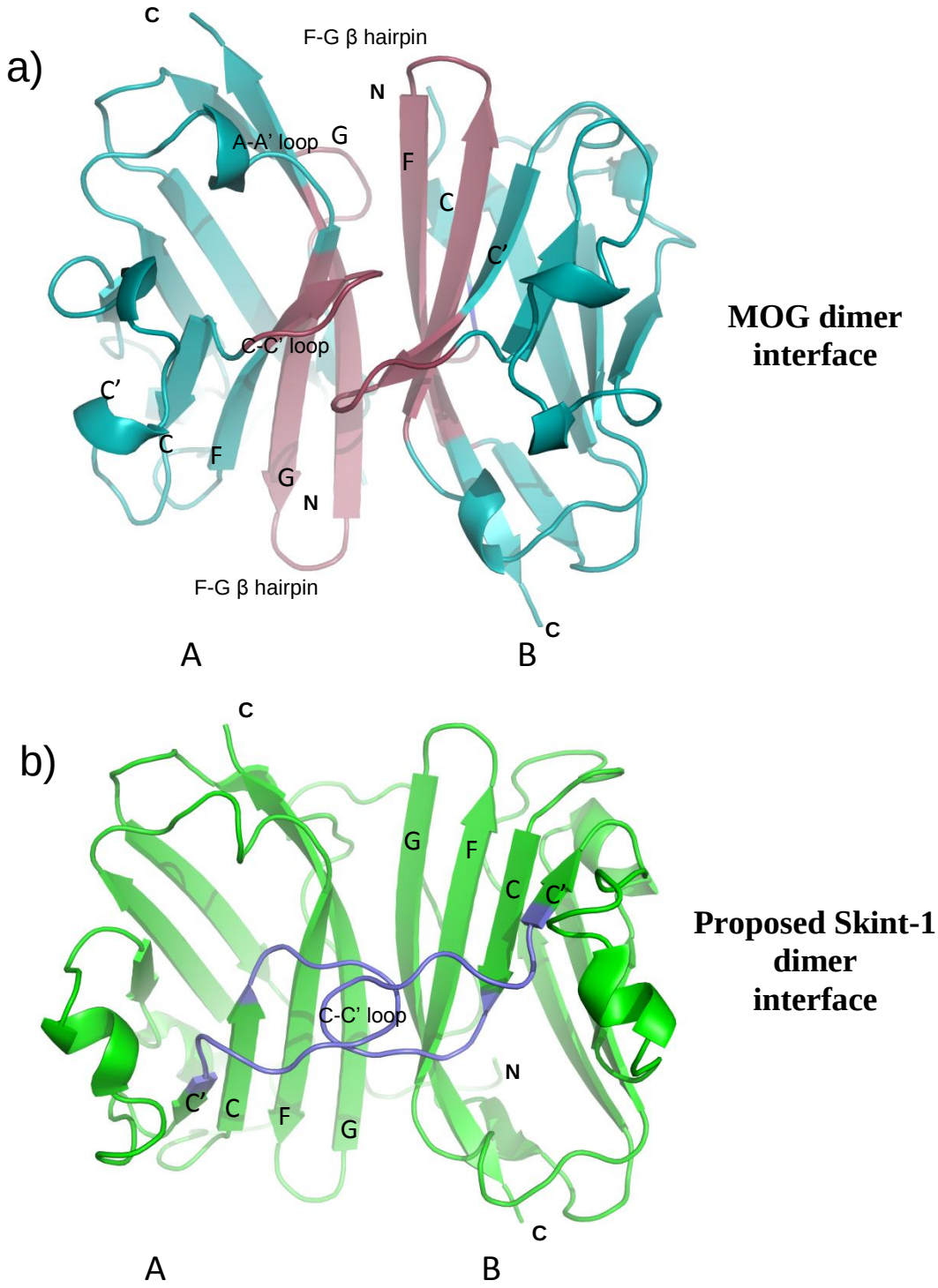


Figure 6.7 a) Interactions at the MOG dimer interface. Secondary structural elements that highlighted in pink are involved in stabilising the dimer interface (Clements *et al*, 2003). Monomers that form the dimer are labelled A and B, respectively. b) Putative Skint-1 dimer interface generated using the MOG dimer as a template. The longer C-C' loop in Skint-1 (depicted in blue) prevents dimer formation presumably due to major steric clashes with the opposing monomer.

contributing to the dimer interface are semi-conservatively altered (S45, T97, Q106 and A110 in MOG is replaced with T48, H100, Y109 and I113 in Skint-1). Moreover, in MOG, G9 stabilises the dimer interface by mediating a hydrogen bonding interaction with E107 of the opposing monomer via its backbone carbonyl group. In contrast, in Skint-1 a non-conservative substitution to E12 is unlikely to interfere with dimer formation as the backbone carbonyl group is predicted to interact with E110. Consequently, there is no obvious reason for why Skint-1 fails to adopt the mode of dimerisation observed for MOG.

To examine this further, a potential model of the Skint-1 dimer was generated based on using the MOG dimer as a template. In MOG, the C-C' loop of each monomer contributes significantly to stabilizing the dimer interface. However, modelling of the Skint-1 IgV dimer demonstrates that the C-C' loop, which is considerably greater in length compared to MOG, is orientated at a perpendicular angle that protrudes out of the β sheet, resulting in major steric clashes with the corresponding C-C' loop of the opposing Skint-1 monomer (Figure 6.7b). As a result, it is predicted that due to steric hindrance, the mode of dimerisation observed for MOG is structurally incompatible with Skint-1 IgV.

Similarly to MOG, the dimer interface of the PD-L1 IgV domains is stabilised predominantly by multiple hydrogen-bonding interactions (Figure 6.8a), albeit with a significantly reduced buried surface area (PD-L1=859Å^2; MOG=1800Å^2). The regions that contribute to stabilising the PD-L1 dimer interface include strand A (T22), strand B (E39 and K41), B-C loop (V44) and α helix 4 (Q91). Superimposition of the Skint-1 IgV domain onto the PD-L1 dimer has enabled the generation of a putative Skint-1 dimer model (Figure 6.8b). Interestingly, the majority of residues highlighted to be important for contributing to stabilising the PD-L1

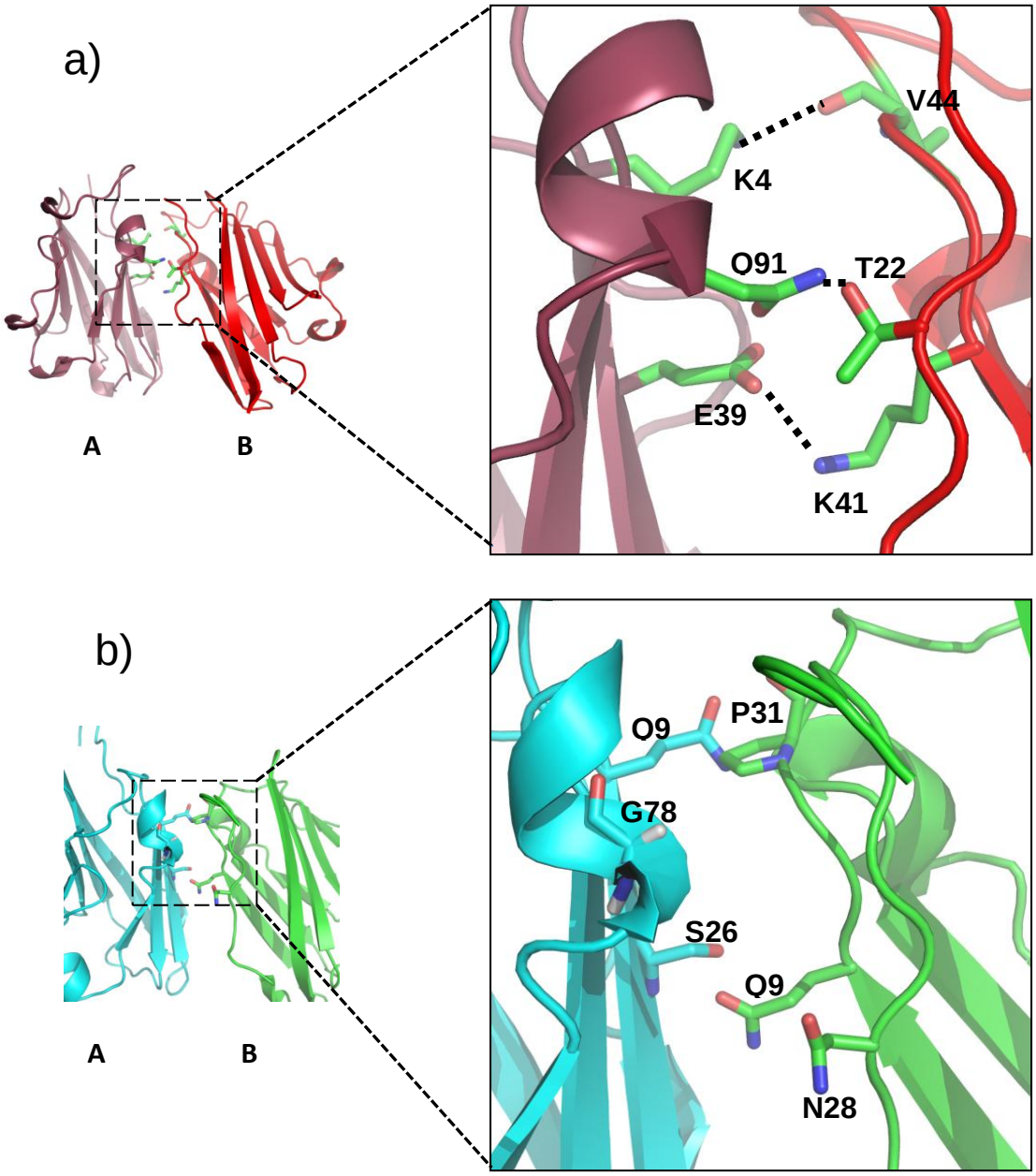


Figure 6.8 a) PD-L1 dimer with a close-up view of the interactions that contribute to stabilising the interface (Chen *et al*, 2010). Monomers A and B are coloured in burgandy and red, respectively. Contacting residues are shown in ball-and-stick format, and hydrogen bonds are represented by black dashed lines. b) Putative Skint-1 dimer interface generated using PD-L1 dimer as template. The close-up view demonstrates clashes and loss of interactions at the proposed Skint-1 dimer interface. Monomers A and B are coloured cyan and green, respectively.

dimer interface are non-conservatively substituted in Skint-1. Consequently, the molecular contacts present at the PD-L1 dimer interface are likely to be absent for Skint-1. For instance, whereas in PD-L1, E39 forms a salt bridge contact to K41 (on the corresponding monomer), non-conservative substitutions to S26 and Q28 would result in loss of this interaction (Figures 6.8a/b). Similarly, whereas in PD-L1 the Q91 side chain hydrogen bonds with the hydroxyl moiety of T22 (opposing Skint-1 monomer), introduction of the G78 and N9 at the equivalent positions in Skint-1 would abolish this interaction (Figures 6.8a/b). Finally, whereas K41 of PD-L1 forms hydrogen bonding interactions with the backbone carbonyl of V44 (on the second PD-L1 monomer), in Skint-1 non-conservative substitutions to Q28 and P31 would result in a loss of this contact as well as introduce major steric clashes (Figures 6.8a/b). Collectively, these observations explain why the Skint-1 IgV domain is unlikely to conform to the dimerisation mode of PD-L1.

6.4 Comparison of Electrostatics surface potential of Skint-1 with MOG

To identify putative ligand binding sites for Skint-1, a molecular surface electrostatic potential map was calculated using DELPHI. Comparison of the molecular electrostatic potential surfaces for Skint-1 IgV and MOG revealed striking dissimilarities in the surface charges. In particular, crystallographic analysis of MOG revealed a striking electronegative strip that runs approximately half the length of the molecule, terminating at a small patch of electropositive charge (R4, R25 and K80) (Figure 6.9). This electronegative strip is dominated by residues from the N-terminal part of strand B (D18, E19 and E21) with minor contributions made from strands D (E74) and G (E116). Notably, this distinctive electrostatic pattern, which has been proposed as a putative ligand binding site for MOG (Clements *et al,* 2003), is not shared in the corresponding regions of Skint-1, owing to several non-conservative amino acid

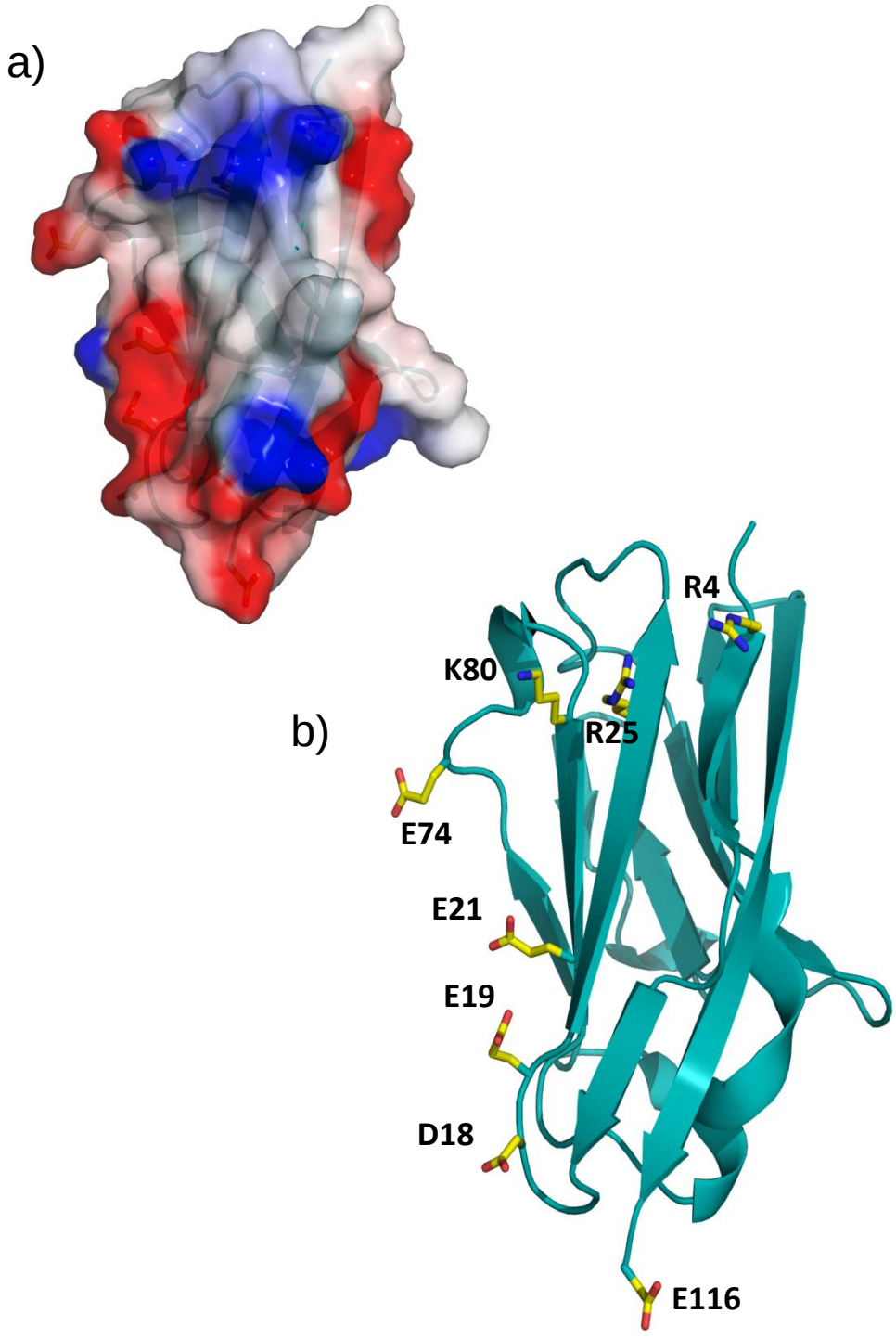


Figure 6.9 Molecular surface representation of the MOG IgV domain. a) Electrostatic potential, highlighting the electronegative strip that terminates at the electropositive patch. b) Residues contributing to these features are labeled and shown in the same orientation as the left panel (Clements *et al*, 2003).

substitutions (Figure 6.10). Further analysis of the Skint-1 IgV molecular electrostatic surface revealed an electronegative strip broadly similar in length to that of MOG, however its location is quite different, positioned on the loop preceding the strand A as well as C, C', G β strands (E4, E39, D59, E110 and E111), and oriented approximately perpendicular to the long axis of the G, F and C strands, relatively close to the membrane-distal region of the domain (Figure 6.11). In addition, Skint-1 IgV exhibits two small electrostatic patches that are absent in MOG. Firstly, an electropositive patch is located slightly above the electronegative strip, generated by R53 and K104 (Figure 6.11). Secondly, an electronegative patch is found in the membrane-distal region of the IgV domain that is generated by the presence of two aspartic residues (D105 and D107). Taken together, these observations suggest that MOG and Skint-1 most likely recognize distinct ligands.

6.4.1 Comparison of Solvent-Exposed Hydrophobic Patches of Skint-1 and MOG

Generally, solvent-accessible hydrophobic surfaces are not energetically favourable and are therefore likely to represent potential protein–protein or protein-ligand binding sites. Indeed, there exists a high correlation between surface exposed hydrophobicity and protein-protein interactions. In this vein, MOG exhibits a number of solvent-exposed hydrophobic residues, including three aromatics: Y10, F44, and F90 (Figure 6.12). The equivalent residues in Skint-1 are semi-conservatively substituted at two of these positions (Y47 and V93). In Skint-1, Y47 and V93 are located at the membrane-proximal region of the Skint-1 molecule and therefore ideally placed to mediate inter-domain interface stabilizing interactions with residues from the IgC domain. Other notable solvent-exposed non-polar side chains located in the membrane-proximal portion of Skint-1 include L19, Y68, V69, F89 and V93, and these may also function in stabilizing the IgV-IgC inter-domain interface. Indeed, since these

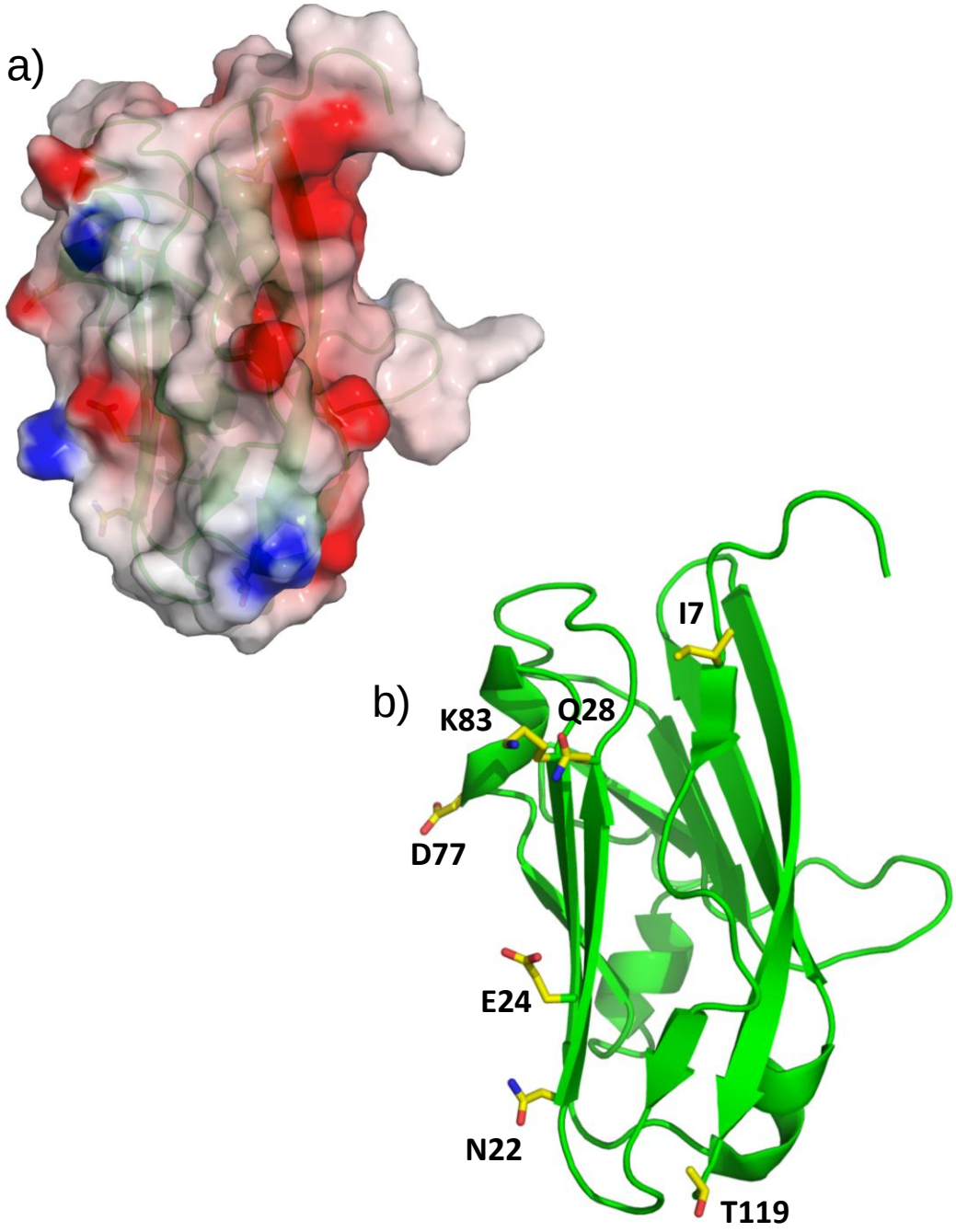


Figure 6.10 Molecular surface representation of the Skint-1 IgV domain. a) Electrostatic potential map highlighting the lack of electronegative strip and electropositive patches near strands B, D and G of MOG. b) Residues contributing to these features are labeled and shown in the same orientation as the top panel.

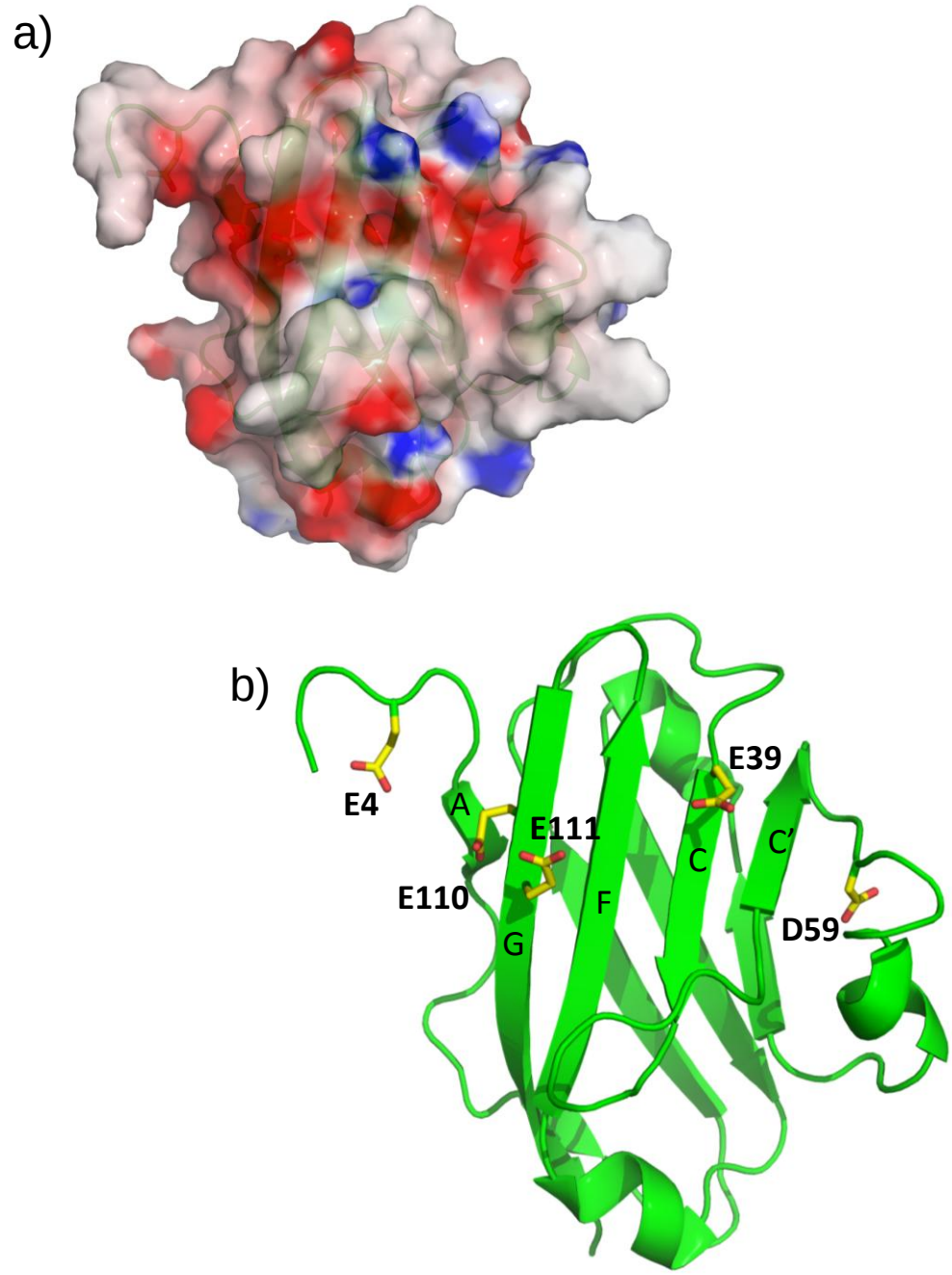


Figure 6.11 Molecular surface representation of the Skint-1 IgV domain. a) Electrostatic potential, highlighting a novel electronegative strip running approximately perpendicular to strands C, F and G. b) Residues contributing to these features are labeled and shown in the same orientation as panel a.

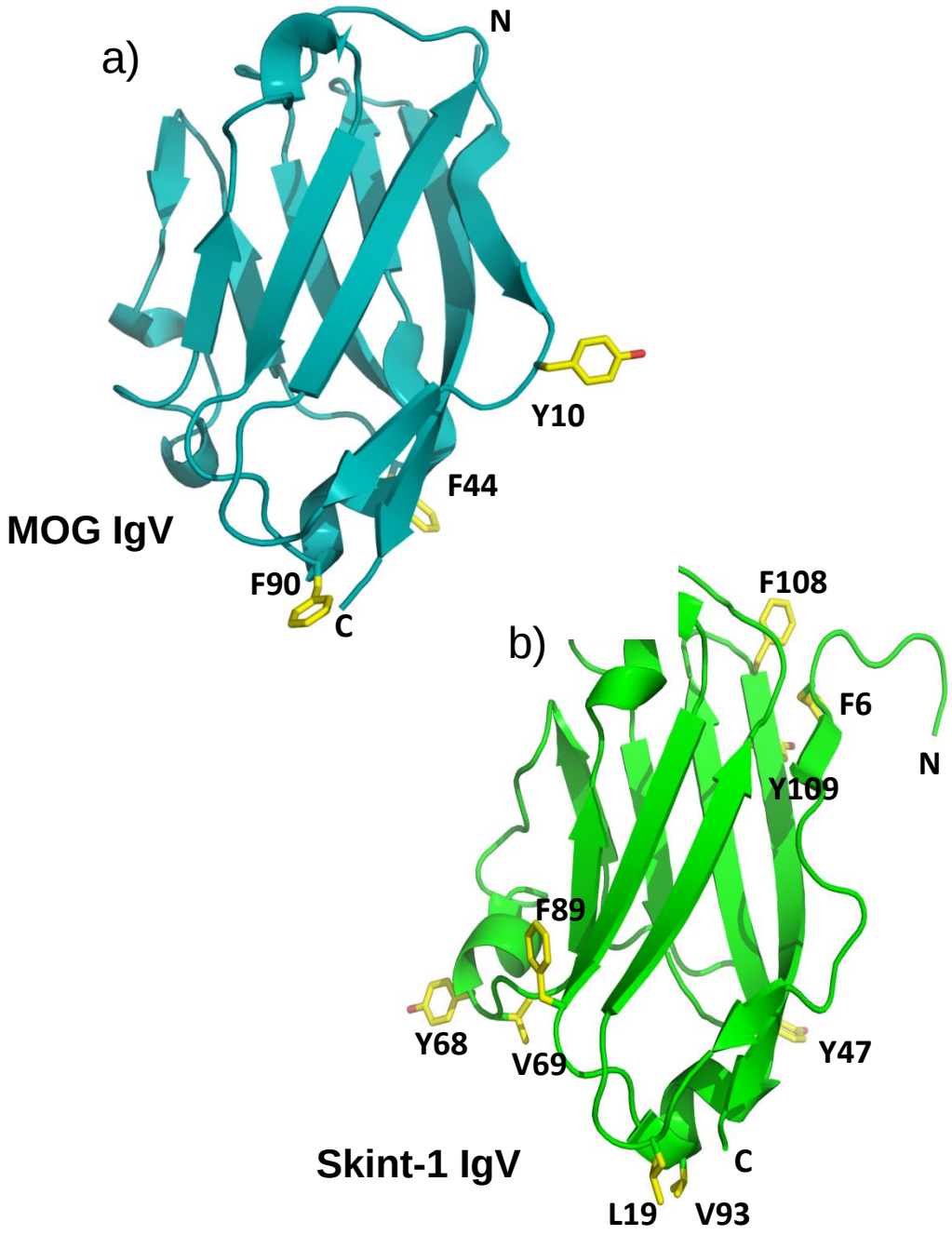


Figure 6.12 a) Putative ligand binding site for MOG. Solvent-exposed aromatic side chains that may be involved in ligand binding are drawn in ball-and-stick format (Clements *et al*, 2003). b) Solvent-exposed non-polar side chains in Skint-1. Residues located at the C-terminal region are most likely involved in stabilising the IgV-IgC interdomain interface.

residues are highly conserved across the Skint gene family, they most likely represent a common structural and functional feature. Strikingly, there are several solvent-exposed aromatic side chains located in the membrane-distal end of Skint-1 IgV (Figure 6.12), which will be discussed in greater detail in the subsequent section.

6.4.2 Characterisation of the CDR-like loops of Skint-1 with MOG

Determination of the Skint-1 IgV structure has allowed detailed characterisation of potentially functionally significant regions of the Skint-1 molecule. In particular, the three loops at the membrane distal end of Skint-1 that connect the B-C (Q28-M38), C-D (D56-I64) and F-G (K104-Y109) strands have been termed CDR1/2/3 like loops, respectively (Figure 6.13). These CDR-like loops are in the equivalent positions to the Complementarity Determining Regions (CDRs 1, 2, and 3, respectively) of IgV domains present in T cell receptor or B cell receptor chains. Intriguingly, at the membrane-distal end of Skint-1, there is a small hydrophobic patch in and around CDR3 like loop, comprised of three solvent exposed aromatic rings (F6, Y108 and Y109) (Figure 6.13). It is possible that the recently identified functionally relevant F-G loop (personal communication with Professor Hayday laboratory) in combination with this trio of solvent-exposed hydrophobes form a potential protein-protein docking site. Interestingly, the equivalent region in MOG is largely conserved with F3, Y105 and Q106, suggesting that this area may also be important for ligand docking (Figure 6.13).

Comparison of the CDR-like loops between Skint-1 and MOG highlights several striking differences in their overall sequence, length and conformation. In MOG, the three loops combine to form a polar and essentially uncharged featureless molecular surface (Figure 6.14). In contrast, in Skint-1 the membrane-distal loops collectively form an undulating and

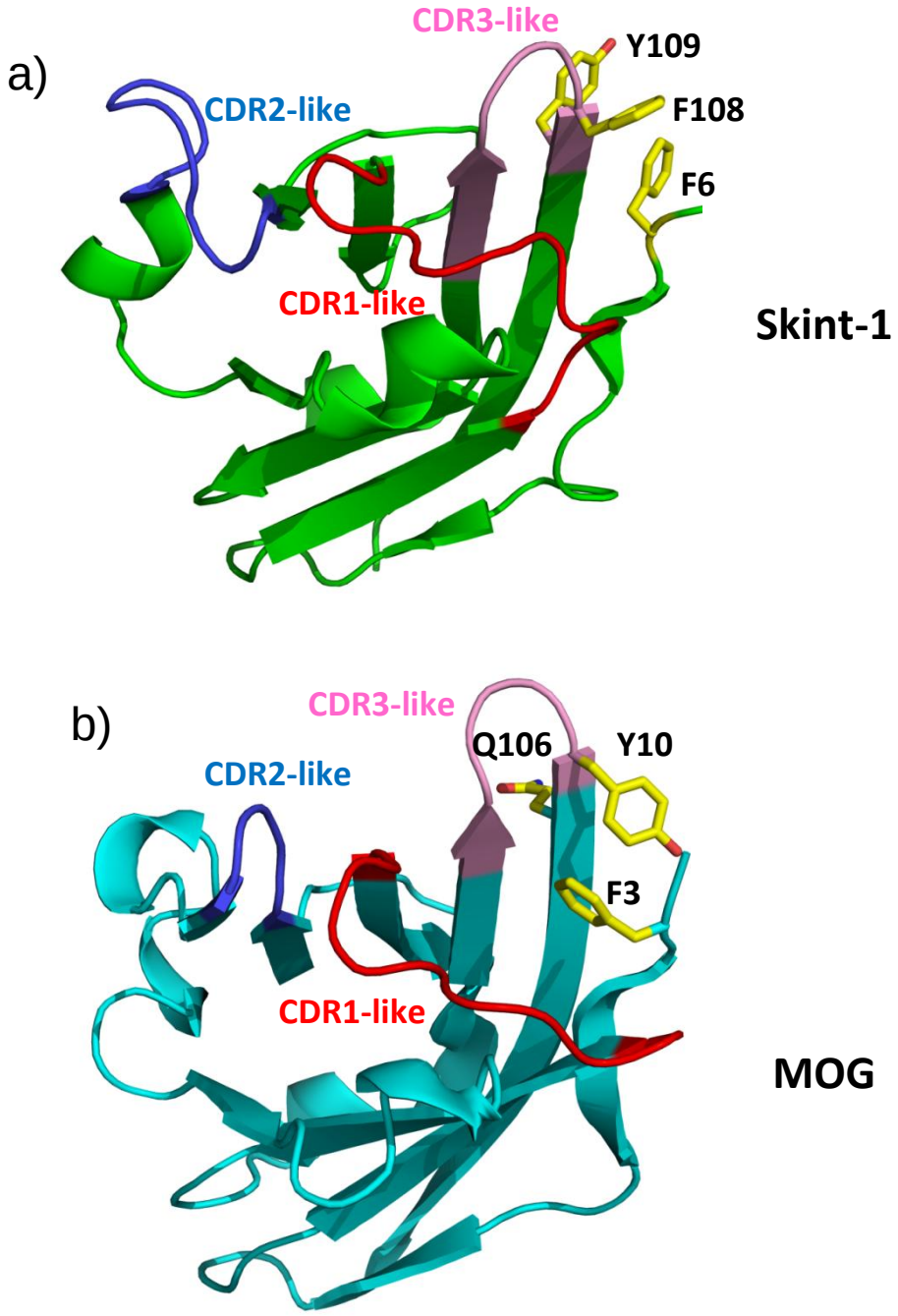


Figure 6.13 Key features of the membrane-distal regions of Skint-1 and MOG. a) Skint-1 loops that connect B-C (CDR1-like), C-D (CDR2-like) and F-G (CDR3-like) strands are coloured in red, blue and pink, respectively. Trio of solvent-exposed aromatic side chains are located in and around the CDR3-like loop. b) MOG loops that connect B-C (CDR1-like), C-D (CDR2-like) and F-G (CDR3-like) strands are highlighted in red, blue and pink, respectively (Clements *et al*, 2003). Note that two of the solvent exposed aromatic residues in Skint-1 are conserved in MOG, indicating a common putative protein binding site.

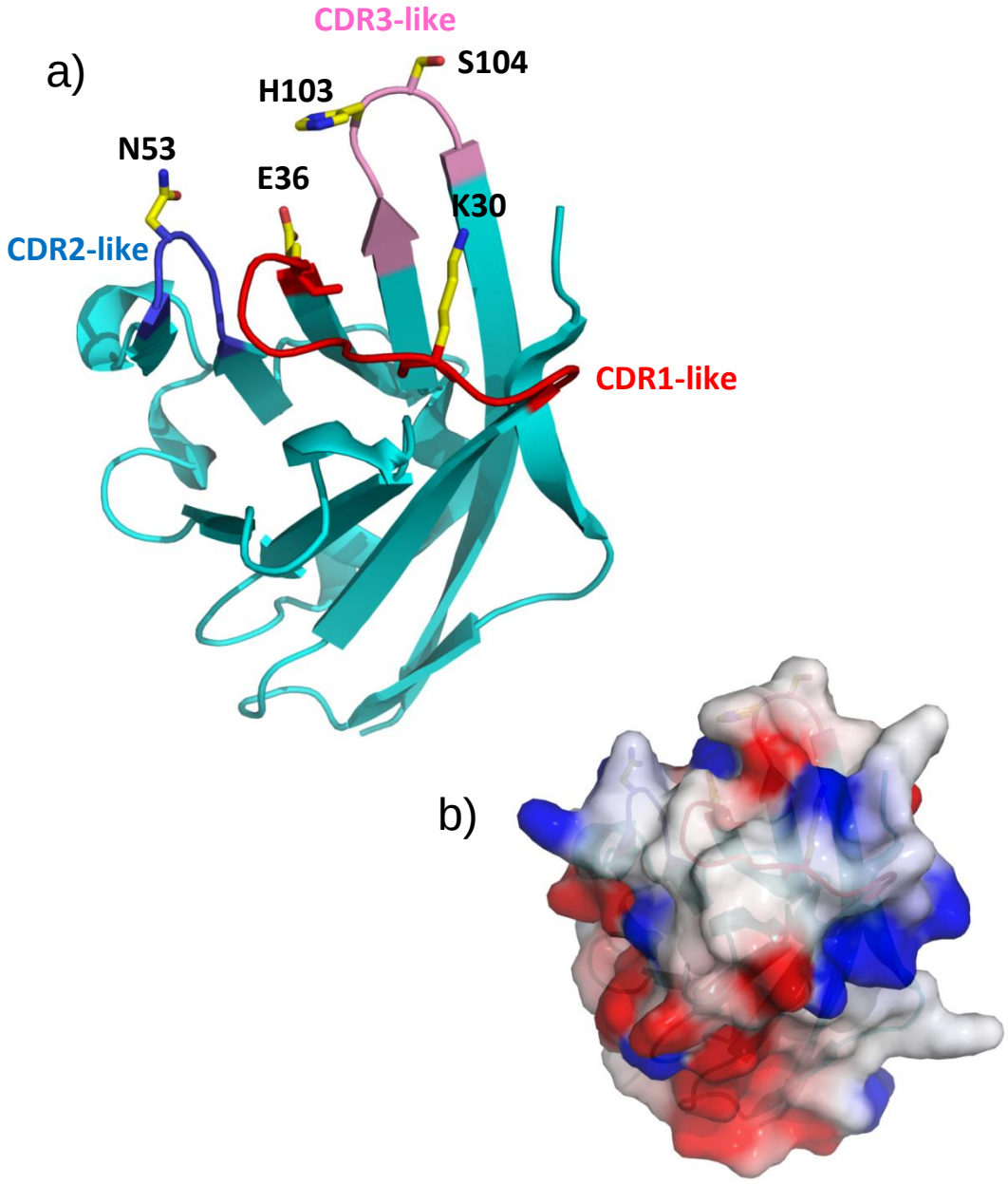


Figure 6.14 Key molecular features associated with MOG CDR-like loops. a) Residues from CDR-like loops that are found in highly prominent positions are highlighted in ball-and-stick format. The second residue in the CDR3-like loop is occupied by H103 which strains the conformation of the type II β-turn. MOG CDR-like loops are highlighted in red, blue and pink, respectively. b) Electrostatic potential map for membrane-distal region of MOG (orientation as in panel a). The three CDR-like loops combine to form essentially a flat and uncharged molecular surface (Clements *et al*, 2003).

charged molecular surface (Figure 6.15). The CDR1-like loop in Skint-1 (28QLSPPQQAQHM38) is relatively similar in length to MOG (27SPGKNATGME36). However, whereas in MOG the B-C loop forms a flat and indistinctive surface, in the corresponding region of Skint-1 the side chains of Q33, Q34 and H37 protrude prominently to form an undulating polar surface (Figure 6.15). In addition, there are significant differences observed between the CDR2-like loop of Skint-1 (56DGKDMFGEI64) and MOG (51YRNG54). Firstly, the CDR-like loop in Skint-1 is considerably longer in length as it lacks the C'' region found in the equivalent region of MOG. Secondly, whereas in MOG the loop is comprised of mainly non-polar and charged residues, in Skint-1 the loop displays amphipathic properties, containing both hydrophobic and hydrophilic residues. Finally, in Skint-1 the side chain moieties of D56 and K58 are highly solvent exposed thereby providing a charged molecular patch that is potentially available for ligand binding (Figure 6.15).

In both homologues, the CDR3-like loop is located in a highly prominent position of the IgV domain, forming a hairpin loop that adopts a type II β-turn. However, there are several distinctive features evident in Skint-1 that are absent in MOG. In MOG, the second residue of the F-G loop is occupied by a non-canonical H103, whereas in Skint-1 a glycine is present, consistent with what is usually expected for a type II β-turn (Sibanda, B. L. & Thornton, J. M.1985). As a result, the dihedral angles of H103 are found in disallowed regions of the Ramachandran plot, allowing MOG to adopt a strained loop conformation. It is well documented that residues with strained conformations are often positioned at functionally important sites in proteins. This coupled with the fact that the side chains of H103 and neighbouring S104 protrude from the tip of CDR3-like loop, suggest that this region represents a putative ligand binding site for MOG (Figure 6.14). Indeed, the crystal structure

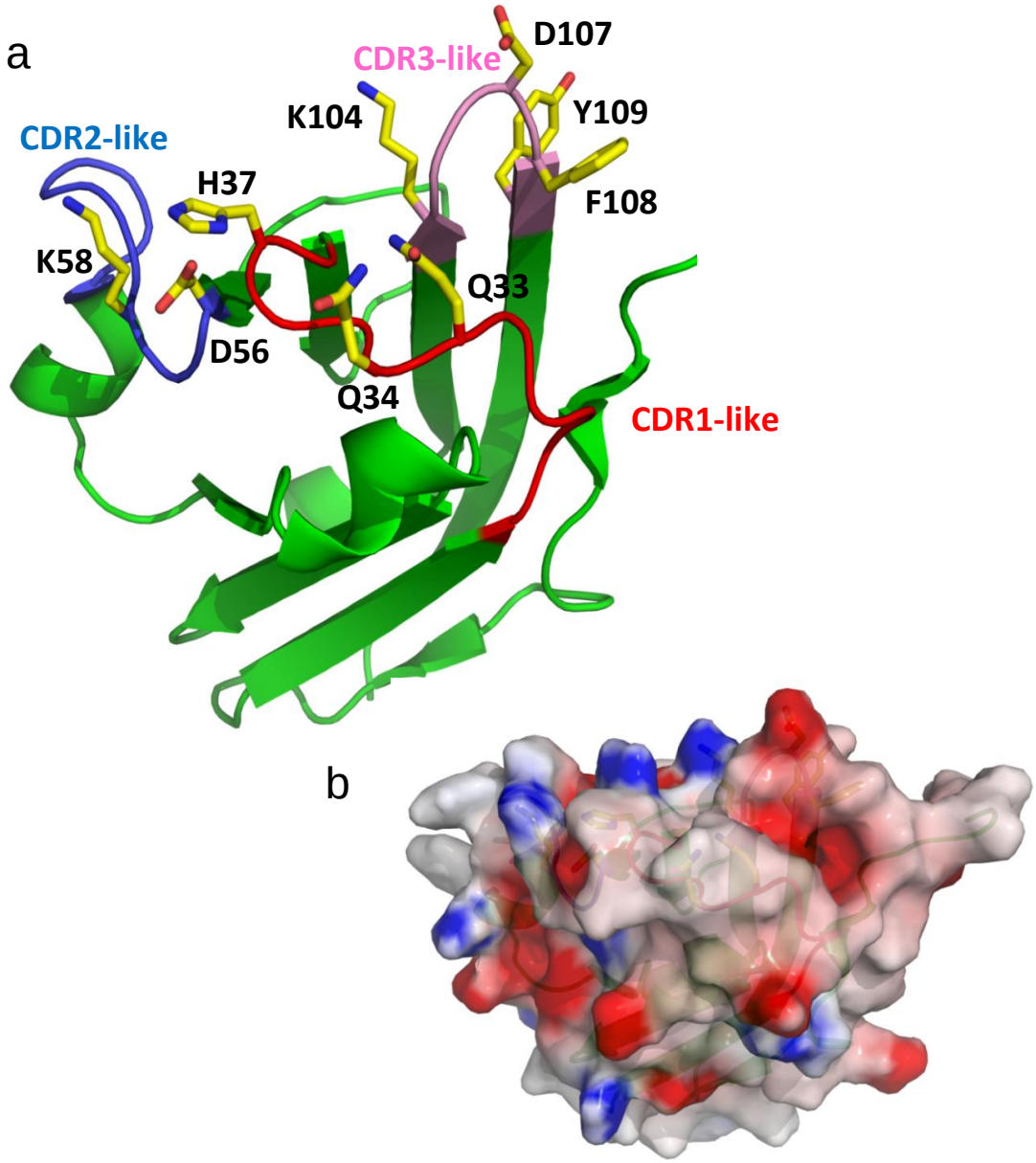


Figure 6.15 Key molecular features associated with Skint-1 CDR-like loops. a) Residues from CDR-like loops that are found in highly prominent positions are highlighted in ball-and-stick format. The second residue in the CDR3-like loop is occupied by canonical G106 which allows the type II β-turn to adopt a more flexible conformation relative to MOG. Skint-1 CDR-like loops are highlighted in red, blue and pink, respectively. b) Electrostatic potential map for membrane-distal region of Skint-1 (orientation as in panel a). The three CDR-like loops combine to form a charged and undulating molecular surface.

of the extracellular domain of MOG in complex with the antigen-binding fragment (Fab) of the MOG-specific demyelinating monoclonal antibody, 8-18C5, has provided further insights into the molecular determinants of MOG ligand binding (Breithaupt *et al*, 2003). The 8-18C5 antibody binds to the three CDR-like loops located at the upper membrane-distal end of MOG with a significant contribution mediated by the G strand and the preceding F-G loop (encompassing residues R101–E108), a region that is unique for MOG (Figure 6.16). In comparison, in Skint-1 the second residue in the type II β-turn is occupied by a canonical glycine, and as a result adopts an unstrained conformation relative to MOG. Furthermore, the CDR3-like loop encompasses several solvent exposed and highly prominent residues that are essentially polar and charged (K104, D107 and Y109) (Figure 6.15). Collectively, based on these structural differences in the CDR-like loop regions this may account for the functional differences between Skint-1 and MOG.

6.5 Comparison of Skint-1 IgV with other Skint gene family members

The IgV domains of the Skint family proteins can be divided into three subfamilies based on the degree of sequence identity to Skint-1 (Figure 6.17). Firstly, Skint family members 2-6 share 74-81% sequence identity to Skint-1, with the highest levels of conservation found with Skint-3 closely followed by Skint-2. In contrast, the second subgroup consists of Skint-7 and 8 which share moderate degrees of sequence conservation with the Skint-1 IgV domain at 43-45% sequence identity. Finally, the subgroup that exhibits the lowest levels of sequence conservation to Skint-1 contains Skint-9 and Skint-10, with sequence identities ranging from 19-20% (alignments not shown).

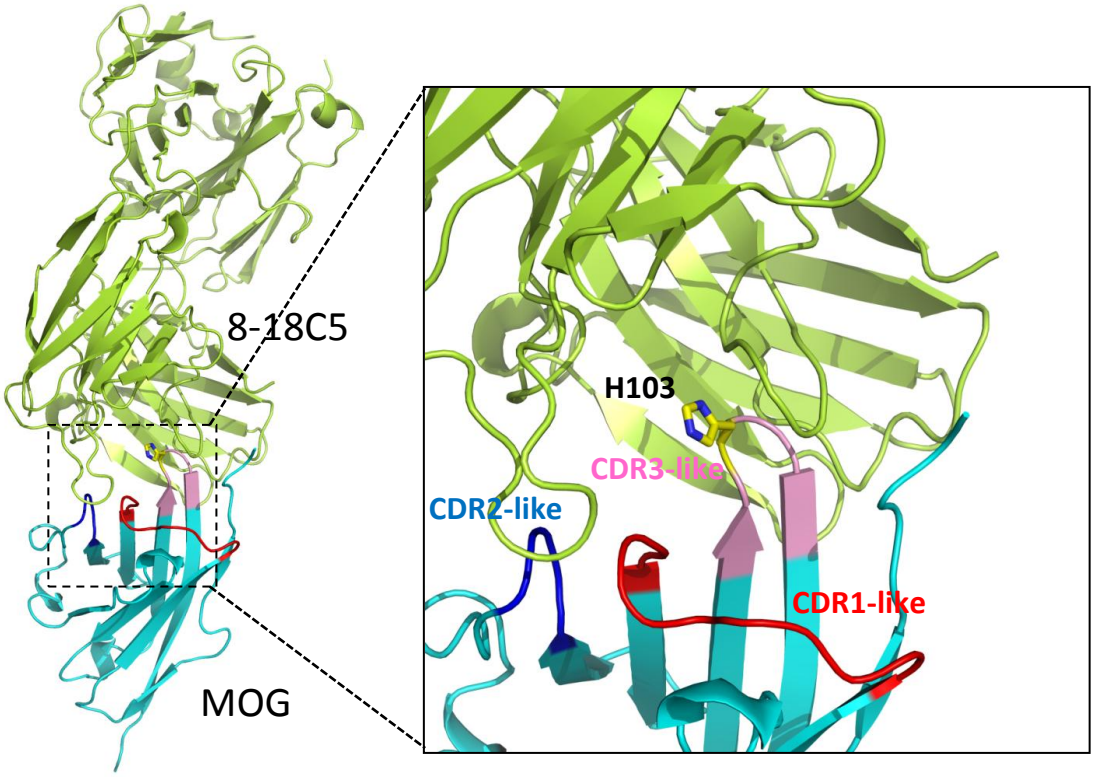


Figure 6.16 Overall structure of MOG in complex with MOG-specific demyelinating monoclonal antibody, 8-18C5 (left; Breithaupt *et al*, 2003). Close up view of the MOG-8-18C5 Fab binding interface (right panel). Majority of the contacts to the Fab heavy chain are mediated by the CDR3-like loop of MOG. CDR-like loops of MOG are highlighted in red, blue and pink, respectively.

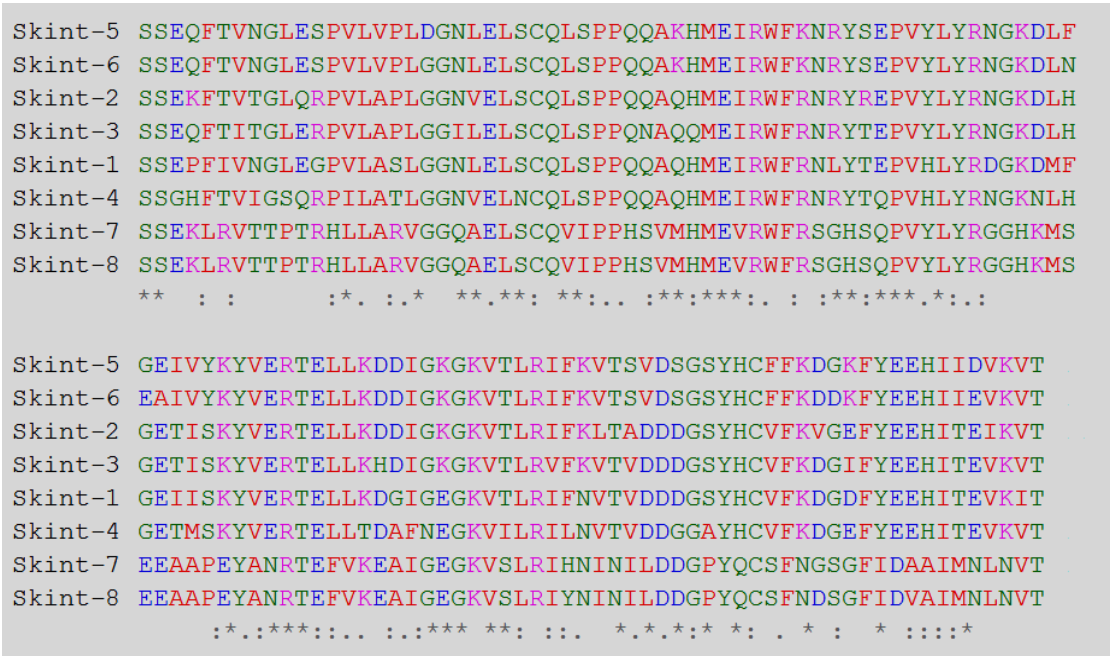


Figure 6.17 Multiple sequence alignment of the IgV domains of Skint gene family. Sequences of IgV domains were downloaded from the Uniprot database and aligned with ClustalW2. *Indicates residues that fully conserved,: represents conservation between groups of strongly similar properties and . Indicates conservation between groups of weakly similar properties. Amino acids shaded in colour based on their chemical properties; red (small and hydrophobic), blue (acidic), magenta (basic) and green (hydroxyl, sulphydryl and amine).

Detailed analyses of aligned sequences for the IgV domains of Skint-1-6 highlighted that many of the regions are highly conserved. In particular, the Hayday group raised a polyclonal antibody against one of these conserved regions, KYVERTELL, located between the first 3_{10} helix and strand D (Figure 6.17). This polyclonal antibody has enabled functional assays to be performed in an effort to dissect the role of Skint-1 in mice DETC maturation. As expected, when mapping the highly conserved segments onto the Skint-1 structure they correspond to core secondary structural elements that contribute to formation and stabilisation of the classical IgV domain. Furthermore, the least conserved regions are restricted to the loops that interconnect the beta strands, particularly the CDR3-like loop, which has recently been identified as playing a pivotal role in the biological activity of Skint-1. The remainder of the structural analysis is limited to characterising the differences between Skint-1, Skint-2 and Skint-7, the only Skint family members to be expressed in mice.

6.5.1 Comparison of Skint-1 Electrostatic molecular surface properties with Skint-2

A model of Skint-2 was generated with PHYRE using the Skint-1 structure as a modelling template. As expected, the high level of sequence identity between Skint-1 and Skint-2 (80%) correlates with an extremely large degree of structural similarity between these mouse Skint gene family members. In total there are 20 amino acid substitutions found in Skint-2, the majority of which, when mapped onto the Skint-1 structure, are located on the surface (Figure 6.18). Notably, nine of these alterations represent non-conservative amino acid substitutions. As a result, the high level of sequence diversity observed in solvent accessible positions presumably reflects the functional specificity of Skint-1. To assess this further, an electrostatic potential map for the Skint-2-PHYRE generated model was calculated with DELPHI. Comparison of the surface electrostatics between Skint-1 and Skint-2 show a

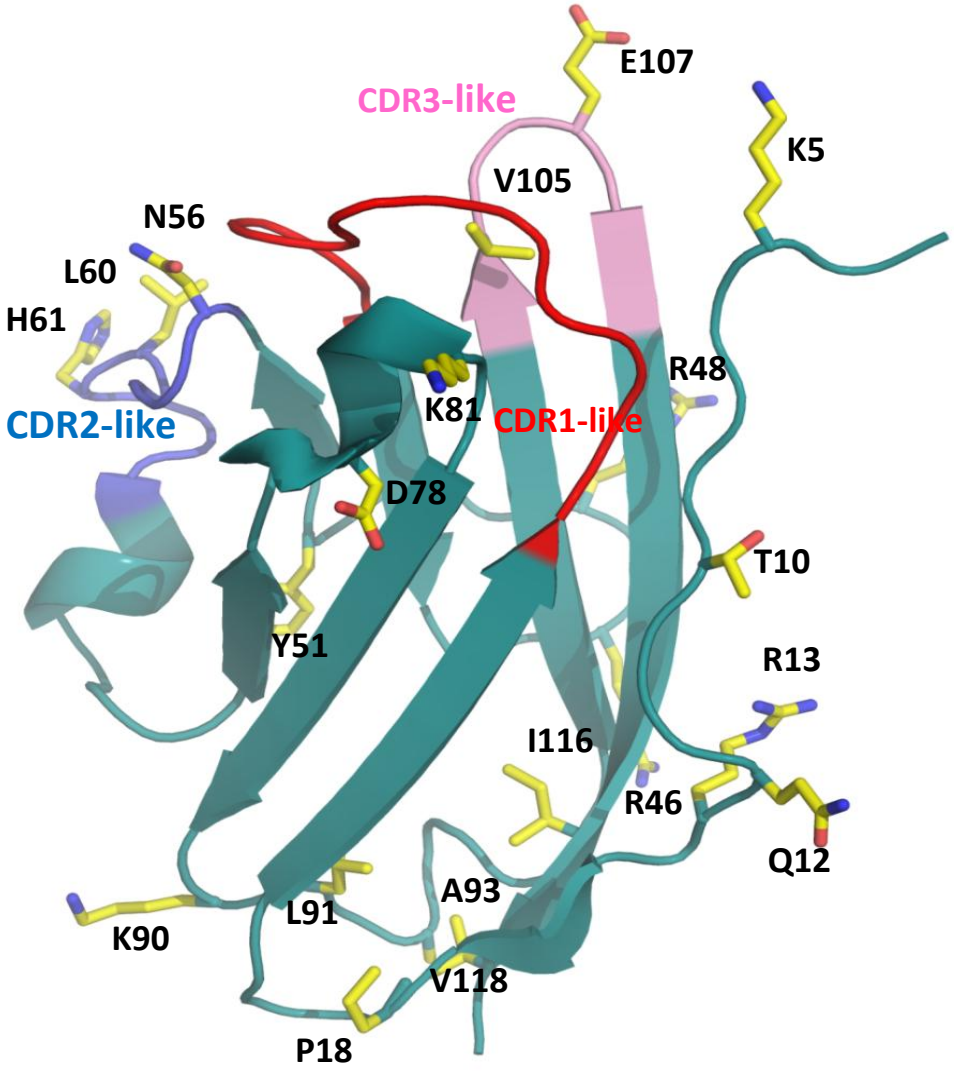


Figure 6.18 Model of Skint-2 IgV domain generated using Phyre. Residues that differ from Skint-1 are highlighted in ball-and-stick format. The Skint-2 CDR-like loops are highlighted in red, blue and pink, respectively.

remarkably similar electronegative strip that runs perpendicular to the long axis of the G, F and C strands, relatively close to the membrane-distal region of the domain. There is a high degree of sequence identity for these negatively charged residues across the Skint family members 1-6, and to a lesser extent Skint-7, suggesting that the electronegative strip is not likely to be critically important in the non-redundant function of Skint-1. Strikingly, of the nine non-conservative amino acid alterations between Skint-1 and Skint-2, six lead to the insertion of a positively charged residue (K5, R13, R46, R48, K81 and K90) which results in Skint-2 exhibiting a higher number of electropositive surface patches (Figure 6.19). It is possible that these patches may represent a potential ligand binding site that is unique to mouse Skint-2, given that only two of these positively charged residues are conserved in Skint-7.

6.5.2 Comparison of Electrostatic molecular surface between Skint-1 and Skint-7

The Skint-1 IgV structure was used as a template to generate a first-order model for the most divergent Skint family member in mouse, Skint-7, using PHYRE. Despite the lower sequence identity observed between Skint-1 and Skint-7 (44%), the overall structure of Skint-1 is very homologous to the PHYRE generated Skint-7 model. Interestingly, similarly to Skint-2, the least conserved residues between Skint-7 and Skint-1 are located in the solvent exposed regions of the protein. The solvent exposed non-conservative amino acid substitutions in Skint-7 have resulted in marked differences in the electrostatic molecular surface when compared to Skint-1. Firstly, Skint-7 does not possess the extensive electronegative zone running perpendicular to the long axis of the G, F and C strands as found in Skint-1 and Skint-2, owing to several amino acid alterations (D59K and E111A) (Figure 6.20). Secondly, there are additional electropositive patches found in Skint-7, relative to Skint-1, due to the

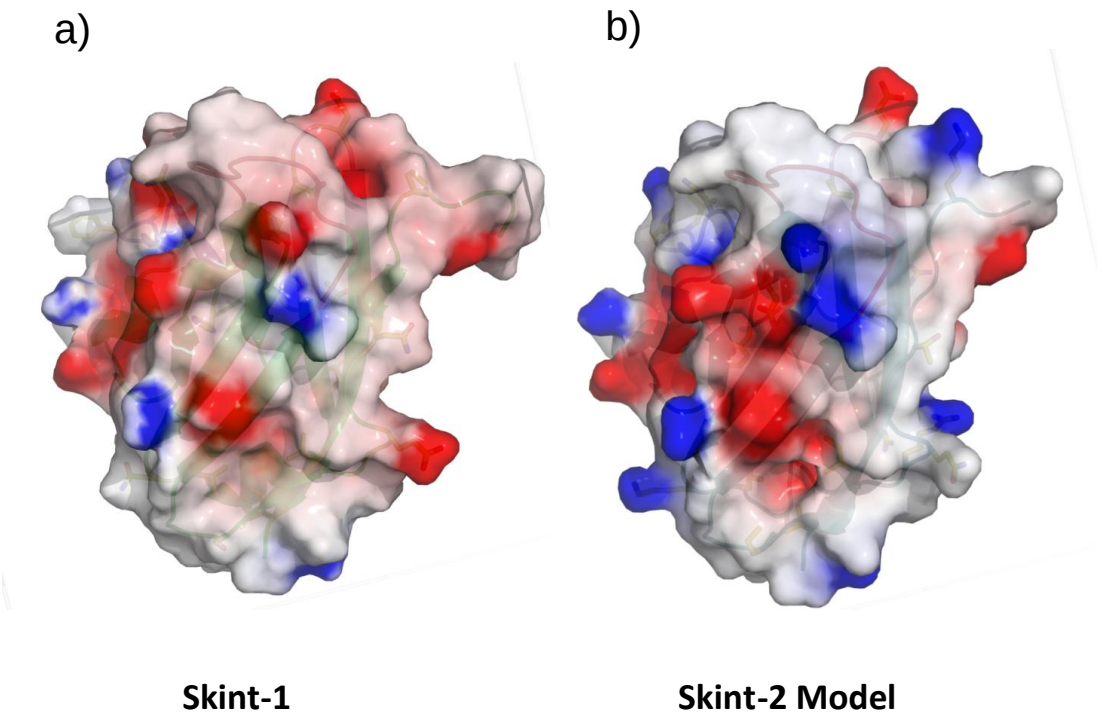


Figure 6.19 Comparison of electrostatic molecular surfaces between the IgV domains of Skint-1 (a) and Skint-2 Phyre generated model (b). Solvent-exposed non-conservative residue substitutions leads to an increased number of electropositive patches in Skint-2 relative to Skint-1. The orientation for each structure is as shown in Figure 6.18.

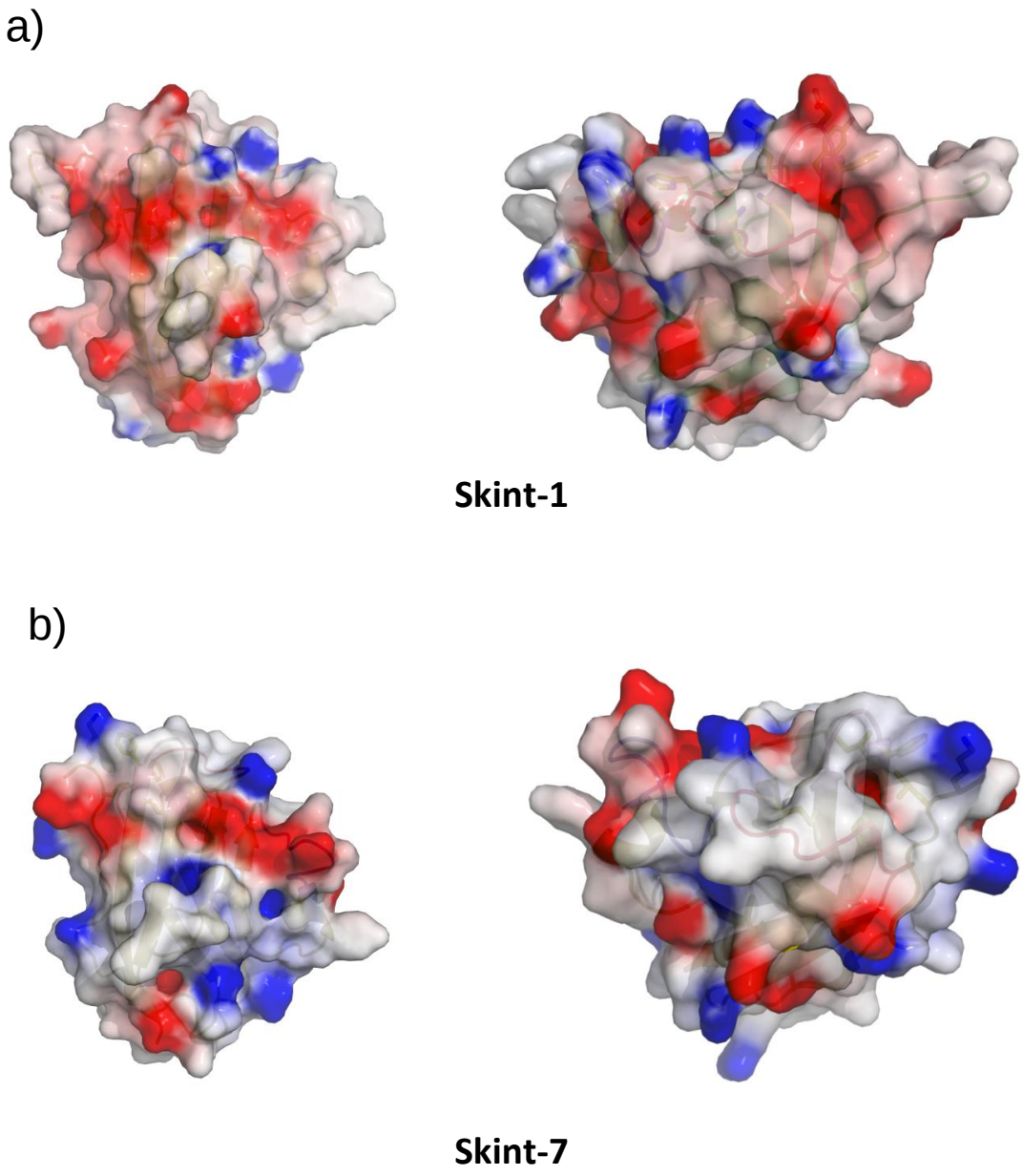


Figure 6.20 Comparison of electrostatic molecular surfaces between the IgV domains of Skint-1 (a) and Skint-7 Phyre generated model (b). The electronegative strip runs discontinuously in Skint-7 (bottom, left) relative to Skint-1 (top, left). The CDR-like loops in Skint-7 (bottom, right) combine to form a featureless and less charged surface relative to Skint-1 (top, right). These changes in the molecular surface may account for differences in ligand binding.

presence of R7, R13 and K56. Finally, the membrane-distal portion of the Ig domain in Skint-7 is essentially uncharged and featureless compared to Skint-1 (Figure 6.20). Given these changes, coupled with the low sequence identity observed between Skint-7 and Skint-1, it provides a possible molecular explanation for the unique role of Skint-1 relative to Skint-7

6.6 Characterization of the CDR-like loops of Skint-1 with Skint-2 and Skint-7

As mentioned previously, the activity of Skint-1 is truly non-redundant as in RTOC experiments involving substitution with other Skint family members, Skint-2 (the closest family member to Skint-1 in mouse) and Skint-7 both failed to rescue DETC progenitor maturation. Moreover, although the Skint-1/Skint2 domain-swap chimeras (involving the IgV, IgC and TM regions) were equivalently expressed as protein, none recapitulated the activity of Skint-1 when expressed in FVB.Tac RTOC, suggesting that each discrete domain of Skint-1 was essential for its physiological function (Barbee *et al*, 2011). More recently, to pinpoint which regions of Skint-1 are biologically important in the positive selection and development of DETC γδ, Skint-1/Skint-2 CDR3-like loop swap chimeras were generated and expressed in FVB.Tac RTOC. The Skint-1 chimeras incorporating CDR1 and CDR2-like loops of Skint-2 had no diminishing affect on the DETC γδ maturation process. However, replacement of CDR3-like loop of Skint-1 (104KDGDFY109) with the equivalent region of Skint-2 (KVGEFY) negated the positive selection of DETC γδ cells (unpublished data; Professor Adrian Hayday's group). Comparison of the CDR3-like loop between Skint-1 and Skint-2 identified single non-conservative (D105 in Skint-1 is replaced by V105 in Skint-2) and conservative (D107 in Skint-1 is replaced by E107 in Skint-2) substitutions that contribute to the unique biological activity of Skint-1. Notably, individual site directed mutants of D105 and D107 with alanine substitutions prevented the development of DETC γδ cells, thereby

confirming the functional significance of each negatively charged moiety. Strikingly, the side chain of D107 is extremely prominently exposed on the membrane-distal end of Skint-1, and therefore may directly interact with an as yet unidentified ligand (Figure 6.21), a process that I hypothesise is essential for the positive selection of DETC γδ cells within the thymus. One possibility is that the selection and development of DETC γδ and subsequent thymic egress and homing to the epidermis may stem from a direct interaction between Skint-1 and the DETC Vγ5Vδ1 TCR. In Skint-2, the corresponding residue at this position is conservatively altered to glutamic acid (E107). Although this change preserves the charge relative to Skint-1, introduction of the longer side chain in Skint-2 may sterically clash with the molecule on the surface of DETC (Figure 6.22). Therefore, steric hindrance presumably perturbs binding to the unidentified ligand on DETCs rather than any discernible changes to the conformation of CDR3-like loop. Furthermore, in Skint-1 Ig, V105 is partially buried in a hydrophobic environment, forming van der Waals interactions with the aliphatic groups of P30, M38 (from CDR2-like loop) and F108 (CDR3-like loop) (Figure 6.21). Hence, it is possible that the introduction of a charged residue at position 105, in what is a predominantly a non-polar environment, affects the conformation of the CDR3-like loop such that it permits binding to the unidentified receptor on DETCs. In comparison, in the equivalent region of Skint-2, a non-conservative amino acid alteration (D/V105) could allow the valine side chain to mediate extensive hydrophobic interactions with nearby non-polar side chains of P30, M38 and F108 (Figure 6.22). Consequently, the more favourable interactions mediated by the valine side chain in Skint-2 may induce the CDR3-like loop to adopt a contrasting conformation to Skint-1, such that it precludes binding to the DETC receptor (Figure 6.22), a hypothesis that could be resolved by determining the three dimensional structure of Skint-2. In comparison, the CDR3-like loop in Skint-7 (104NGSGFI109) possesses two non-conservative modifications

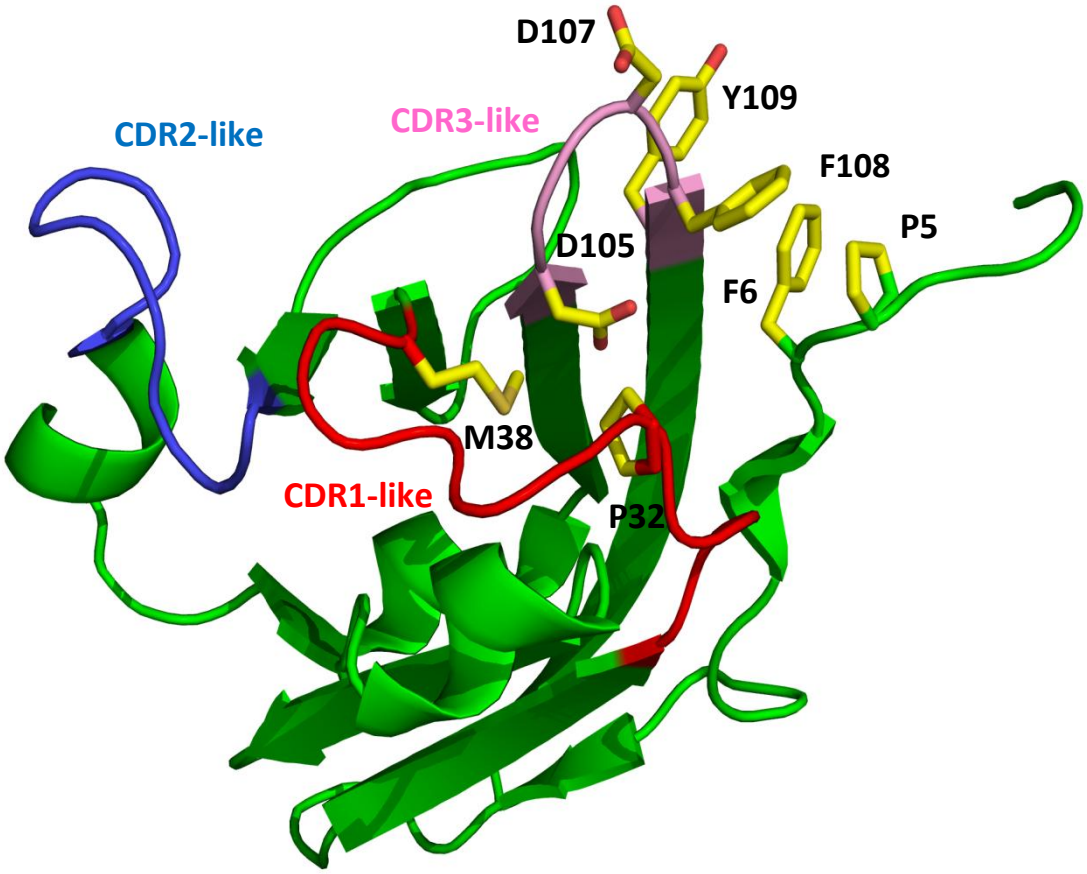


Figure 6.21 Putative Skint-1 binding site for the DETC selecting ligand. Mutations to D105 and D107 prevent the selection and development of DETC. D107 is extremely prominently located suggesting a direct interaction with DETC selecting ligand. D105 is found in an unusual hydrophobic environment (interacting with CDR2-like loop residues P32 and M38) which may allow the CDR3-like loop to adopt a strained conformation such that it permits binding to the DETC selecting ligand. Also, residue P5, unique among Skint gene family, may change the direction of main chain away from the putative binding site permitting ligand binding. Residues that are potentially important in ligand binding are drawn as ball-and-stick format. The CDR-like loops are highlighted in red, blue and pink, respectively.

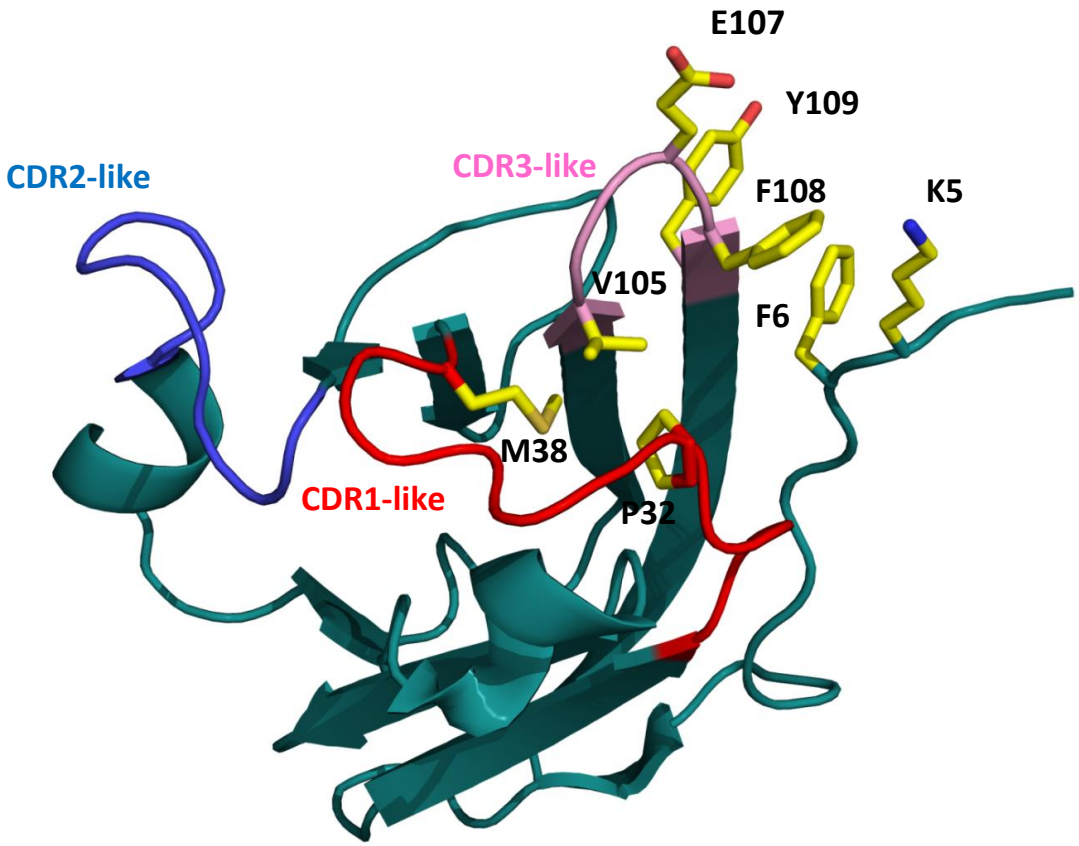


Figure 6.22 Membrane-distal regions of Skint-2. Skint-1/Skint-2 CDR3-like loop swap chimeras prevent the selection and development of DETC. There are 3 possible mechanisms for Skint-2 failing to reconstitute DETC in FVB.Tac RTOC. Firstly, the longer E107 side chain may directly interfere Skint-2 ligand binding. Secondly, V105 mediates multiple hydrophobic interactions with CDR2-like loop residues (M32 and P38) which may affect the conformation of the CDR3-like loop, such that it prevents binding to the DETC receptor. Finally, K5 at the N-terminus may also interfere with ligand binding. Residues are represented as ball-and-stick format. CDR-like loops are highlighted in red, blue and pink, respectively.

at positions that are critically important for Skint-1 activity. In both instances, the negatively charged residues are replaced with glycine (D/G105 and D/G107) and as a result Skint-7 is likely to present a highly flexible and featureless CDR3-like loop which may perturb binding to the DETC selecting ligand (Figure 6.23). Finally, in Skint-2 and Skint-7, there is a positively charged residue (K5) that protrudes from the N-terminus and is located near the vicinity of the CDR3-like loop and the trio of solvent-exposed hydrophobes (distance of ~7Å) (Figures 6.22-23). It is postulated that this long and positively charged side-chain may also perturb ligand binding in Skint-2 and Skint-7. In contrast, in Skint-1 the corresponding region contains a Proline residue which is less prominent and unlikely to interfere with ligand binding. In addition, the Proline side chain reverses the direction of the Skint-1 main-chain away from the putative ligand binding site (Figure 6.21). Intriguingly, given that P5 is not conserved across the Skint gene family members (Figure 6.17), it provides additional circumstantial evidence in support of this residue being important for contributing to the non-redundant role of Skint-1. Collectively, these structural observations provide three putative molecular mechanisms that may ultimately prove critical in the non-redundant role of Skint-1 in facilitating the positive selection and development of DETC γδ cells.

6.7 Identification of Skint-1 Ligand Binding Sites using *de novo* prediction servers

Using the SPPIDER server attempts were made to predict putative Skint-1 IgV protein-protein binding sites. The prediction of potential binding sites was based on integrating solvent accessibility predictions with high resolution Skint-1 IgV structural data. Intriguingly, the *in silico* based predictions highlighted several regions, predominantly at the membrane-proximal end of the Skint-1 IgV module, as possible sites for protein-protein binding.

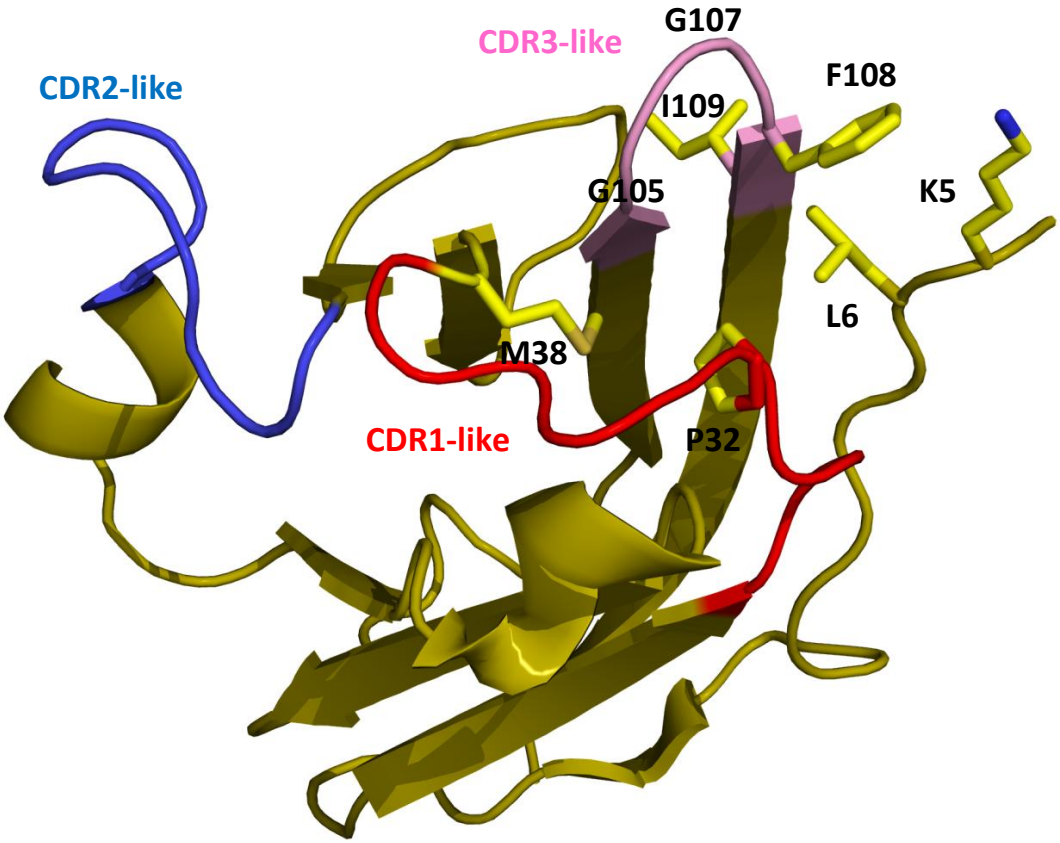


Figure 6.23 Membrane-distal regions of Skint-7. Skint-1 residues implicated in DETC selecting ligand binding (D105 and E107) are altered in Skint-7 (G105 and G107). It is envisaged that such changes provide a featureless and more flexible CDR3-like loop conformation relative to Skint-1 and prevents ligand binding. Also, K5 at the N-terminus may also interfere with ligand binding. Membrane distal residues are represented as ball-and-stick format. CDR-like loops are highlighted in red, blue and pink, respectively.

However, the location of these putative ligand binding residues suggests that the software program has predicted residues that may contribute to stabilising the inter-domain IgV-IgC interface (Figure 6.24). Interestingly, Barbee *et al* highlighted the initial ten amino acids of the IgC domain from Skint-1 as playing a role in selecting DETCs, and these results may add credence to the SPPIDER predicted residues mapped onto the IgV domain. Further attempts with prediction servers ODA and PINUP failed to identify any novel protein-ligand interaction sites, highlighting the difficulties associated with the prediction of protein-protein docking sites.

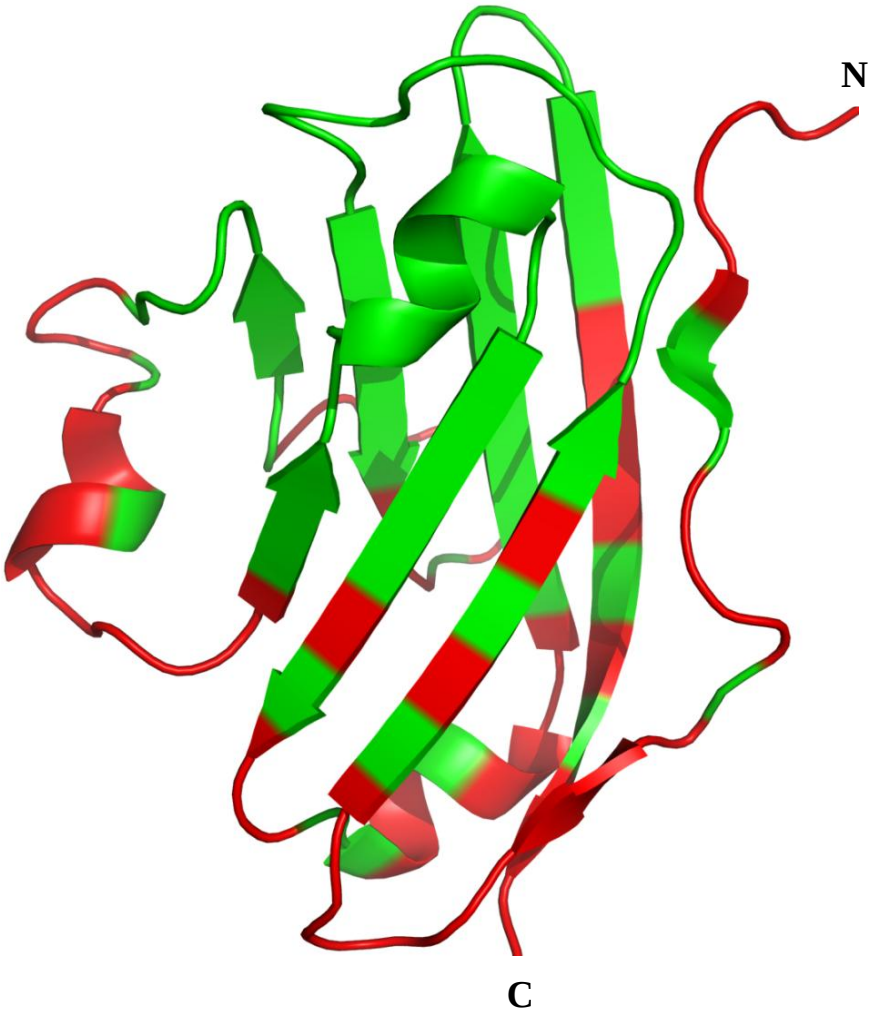


Figure 6.24 Putative ligand binding sites for Skint-1 IgV calculated by SPIDDER. The regions highlighted in red represent potential residues implicated in protein-protein interactions. The majority of residues found in the membrane-proximal region most likely represent regions involved in stabilising the Skint-1 IgV-IgC interface.

6.8 Discussion

The molecular interactions regulating the selection and maturation of the DETC subset, which expresses the Vγ5Vδ1 TCR, are poorly defined. However, substantial evidence now suggests that DETC development is linked to a molecule expressed by thymic stroma, identified as Skint-1 (Lewis *et al*, 2006). Here, the solution structure of the Skint-1 IgV domain has been determined allowing detailed analysis with the ultimate aim of establishing an in-depth molecular understanding of the non-redundant role of Skint-1 in the positive selection of DETCs.

Initial structural determination studies of the Skint-1 IgV domain using X-ray crystallography proved to be unsuccessful due to a combination of issues. Firstly, Skint-1 IgV crystals diffracted X-rays poorly, despite exhibiting excellent morphology and size. A second problem associated with resolving the crystal structure of Skint-1 is the lack of molecular replacement derived phase information. This was unexpected given that the phasing model consisted of MOG, which, similar to Skint-1 secondary structure predictions, adopts an Ig fold and shares significant levels of sequence similarity to Skint-1. The NMR-derived Skint-1 IgV structure has provided insights into the problems encountered during attempts at resolving the crystal structure of the DETC selecting molecule. Initial analysis of the Skint-1 IgV domain revealed insufficient chemical shift data for two distinct regions (S1 to A17 and R44 to F63), suggesting that these areas are extremely flexible. It is possible that, during the crystallisation process, these dynamic regions may have prevented close packing of the protein molecules and hence gave rise to Skint-1 IgV crystals that diffracted X-rays poorly and provided low resolution data (4Å). Comparison of the NMR derived Skint-1 IgV domain structure with the crystal structure of MOG indicated a possible explanation into the failure of molecular

replacement to determine initial phase information whilst using MOG as the phasing model. Indeed, in addition to the two flexible regions (S1-A17 and R44-F63), substantial differences in the conformation of several loops as well as secondary structural elements were observed, which may have interfered with the superimposition process and hence prevented structure determination using molecular replacement.

One of the more remarkable features associated with proteins containing Ig domains is their significant variation in function, despite these proteins exhibiting a similar overall tertiary fold. Indeed, this also appears to be the case for Skint-1 when compared with its closest structural homologue, MOG. These comparisons have revealed several distinct structural features consistent with the fact that the biological role of MOG differs from Skint-1. Firstly, the crystal structure of MOG reveals a physiologically relevant dimeric species, consistent with previous biochemical evidence (Clements *et al*, 2003). In contrast, for Skint-1 the extension of C-C' loop prevents the mode of dimerisation adopted by MOG, confirming that Skint-1 IgV domain exists as a monomeric form in solution as highlighted by the size exclusion profiles during purification. Secondly, comparison of surface electrostatic patterns of the IgV modules of Skint-1 and MOG highlighted several regions encompassing distinctive electronegative patches. Finally, in MOG the CDR-like loop regions combine to form a flat and uncharged surface, whereas an undulating and charged surface is formed in the equivalent region of Skint-1. Taken together, these diverse structural features reflect the unique biological function of Skint-1 relative to MOG. However, these observations are subject to one noteworthy caveat in that the MOG structure was resolved from a crystal structure (Clements *et al*, 2003) whereas the Skint-1 structure was derived from solution NMR data.

Recent data has indicated that Skint-1 plays a unique role in the positive selection and maturation of DETC progenitors. This is somewhat surprising since Skint-1 belongs to a family of genes of which there are currently 10 members that have been identified in the mouse genome. In addition, sequence alignment comparisons have revealed that a few of these family members bear a high degree (~80%) of sequence homology to Skint-1, namely Skint-2 and Skint-3. Initial experiments to assess whether Skint-2 or Skint-7 could substitute for Skint-1 activity in FVB.Tac RTOC revealed that neither family member could restore DETC progenitor maturation. Hence, Skint-1 has a unique and non-redundant role relative to Skint-2 and Skint-7. To determine regions that are essential for the non-redundant biological function of Skint-1 in mice, *in vitro* experiments involving domain swapping chimeras using Skint-1 and Skint-2 were performed (Barbee *et al*, 2011). Skint-2 was used in the domain swap chimeras because this gene is highly homologous to Skint-1 (~80%). These functional studies confirmed that each domain of Skint-1 was indispensable for conferring the maturation of the DETC subset. Further functional experiments in FVB.Tac RTOC with Skint-1 and Skint-2 chimeras highlighted that the amino terminal ten residues of the Skint-1 IgC domain played a pivotal role in DETC progenitor maturation. (Barbee *et al*, 2011).

To elucidate the molecular basis for a differential mode of ligand recognition between the Skint gene family members, I generated structural models of Skint-2 and Skint-7 using the solution structure of Skint-1 IgV as a template. Interestingly, the majority of the non-conserved residues between these molecules were restricted to the surface of the IgV molecule. In addition, Skint-2 and Skint-7 exhibited a series of distinctive electropositive patches that were absent in Skint-1. Collectively, based on these observations it was

envisaged that Skint-2 and Skint-7 are not likely to bind the same ligands as Skint1 and hence perform distinct functions.

More recently, an exciting prospect has involved attempting to further pinpoint the Skint-1 ligand binding site. Indeed, functional assessment of Skint-1/Skint-2 CDR3-like loop swap chimeras in FVB.Tac RTOC have unequivocally demonstrated that the CDR3-like loop of Skint-1 is essential for conferring DETC selection (Personal communication Martin Woodward, Hayday laboratory). Notably, the CDR3-like regions of Skint-1 and Skint-2 are highly conserved with the exception of two residues at positions 105 and 107. In Skint-1, a pair of aspartic acid residues occupies positions 105 and 107, whereas in the corresponding region of Skint-2 there is a valine and glutamic acid. The solution structure of the Skint-1 IgV domain has shed light on the proposed role of D107 in conferring the unique biological function of Skint-1. Intriguingly, D107 in Skint-1 is highly prominent and solvent exposed, strongly suggestive of direct contact with its counter-receptor on the surface of DETCs. In contrast, the equivalent residue in Skint-2 is occupied by a longer side chain, E107, which presumably disrupts binding to a membrane bound receptor on the surface of DETCs. Based on my results, I therefore hypothesise this change in residue at position 107 could impact receptor docking. In addition, since Skint-7 encompasses two glycine residues at these critical positions, it is envisaged that the membrane-distal end of the IgV domain portrays a featureless surface to the counter receptor expressed on DETCs. Interestingly, there are several solvent exposed aromatic hydrophobic residues (F6, Y108 and Y109) in close proximity to the functionally significant CDR3-like region in Skint-1. Given that in most proteins hydrophobic residues are buried within the molecule, it raises the possibility that this group of solvent-accessible hydrophobes are functionally relevant. It is envisaged that these

non-polar aromatics together with the CDR3-like region could form a unique ligand binding site on Skint-1, as they are only present on Skint-1. To systematically evaluate whether any of these potential ligand binding residues are important in the selection of γδ DETCs, a series of Skint-1 variants each incorporating a specific point mutation could be generated and functionally assessed in FVB.Tac RTOC. This analysis would identify and define the limits of such functionally important 'hot spot' regions, and in doing so provide further insights into how Skint-1 recognition of an unidentified receptor(s) leads to the positive selection of DETCs.

Another potentially striking structural feature of Skint-1 is the location and role of P5. This residue, which is present in the loop preceding the A strand, is in close proximity to the CDR3-like loop and the trio of solvent-exposed aromatic hydrophobes. It is hypothesised that the P5 in Skint-1 plays a critical structural role by changing the direction of the main chain away from the proposed ligand binding site. Notably, since the Proline residue is non-conservatively substituted in the remaining members of the Skint family mouse genome, it provides a possible second mechanism underlying the non-redundant functionality of Skint-1. Hence, the lack of P5 in Skint-2 may allow the N-terminal loop to adopt a conformation such that it perturbs binding to the Skint-1 ligand binding partner on DETCs. To address the functional significance of P5 in Skint-1, there are several lines of investigation that could be undertaken. One possible approach is to generate a Skint-1 mutant construct at P5 and assess whether this mutation has an impact on the maturation of DETCs in FVB.Tac RTOC. However, it is important to consider that the generation of a construct encompassing a point mutation of this nature may alter the mechanical stability of the immunoglobulin module resulting in partial or total mis-folding of the protein. An alternative strategy is to determine

the solution structure of the Skint-2 IgV domain and compare the conformation of the loop preceding the A strand with the equivalent region of Skint-1. This type of approach would unequivocally address whether there are significant structural deviations within this region in other family members. In either event, these experiments would highlight whether P5 in Skint-1 promotes a distinct biological conformation, with obvious relevance for its role in the positive selection and maturation of DETCs.

Unfortunately, it was not possible to validate the proposed ligand binding site, identified from the solution structure of IgV module, using several *in silico* protein prediction computational programmes. This underlines the overwhelming challenge in identifying ligand binding sites based solely on *in silico* prediction methods, despite the availability of a plethora of web-based software programs. Nevertheless, the prediction program SPPIDER identified several additional regions as putative ligand binding sites. However, the majority of these were located close to the membrane-proximal region and most likely involved in stabilising the IgV-IgC domain interface. Intriguingly, previous functional experiments with FVB.Tac RTOC have demonstrated that the N-terminal amino acids of IgC are essential for DETC selection. With this in mind, it is conceivable that the regions identified by SPPIDER are critical for binding to the amino terminal 10 residue motif on the IgC domain.

In summary, the solution structure of Skint-1 has identified a putative ligand binding site relevant to its function, however the receptor on the DETC surface that binds Skint-1 is yet to be defined. Strikingly, an obvious candidate that is repeatedly suggested is the Vγ5Vδ1 TCR expressed on the DETC subset. However, at present, this notion is unsupported by any direct evidence. To evaluate this hypothesis, direct protein-protein binding assays could be

performed using surface plasmon resonance or alpha screen technology. The alpha screen is a bead-based assay that has proven to be particularly useful for detection of weak bimolecular interaction between molecules. Although the question of whether the TCR is the receptor for Skint-1 remains ambiguous, the target molecule of Skint-1 is most likely unique to progenitor DETCs (Barbee *et al*, 2011).

CHAPTER 7 –

OVERALL DISCUSSION

Discussion

The aim of this thesis was to carry out studies on both human Vδ2-negative TCRs, and also on the mouse candidate γδ TCR ligand Skint-1 molecule, in order to gain understanding on the molecular basis of γδ T cell recognition. In particular, the key human TCR that I focused on (LES, using Vγ4Vδ5) was derived from a T cell clone that expanded to high levels in a CMV-infected lung transplant patient. In Chapter 3, I successfully produced human recombinant γδ TCRs in an insect cell expression system. In contrast to attempts in *E. coli*, expression of the proteins as TCR-zipper proteins in *Drosophila* was highly successful, resulting in high quality recombinant receptor that was correctly folded. This step was crucial in enabling downstream BIAcore approaches to verify interaction with EPCR ligand, and also to determine the ligand binding mode. In addition, the yield of LES γδ receptor from the *Drosophila* system (~2mg/litre) was substantially higher than in previous work using mouse γδ TCRs. Although the reason for this is unclear, broadly similar yields were obtained for a second human Vδ2-negative receptor from the MAU clone, and this should in principle facilitate structural studies that can require large quantities of protein. Overall, it seems the *Drosophila* expression system is therefore likely to prove a highly useful system for obtaining recombinant TCRs to test interaction with candidate ligands, determine binding modes, and also potentially structural studies.

Structural studies of γδ TCR/ligand interactions are a major aim in the field, and could establish the regions of the γ and δ chains that are involved in ligand binding (e.g. CDR1/2 versus CDR3), which conceivably could help predict which portion of the γδ T cell repertoire is able to recognize specific ligands. My attempts to produce crystals of either the LES γδ TCR in isolation or in complex with EPCR, although unsuccessful, confirmed the yields of

protein recovered for each protein from the *Drosophila* expression system were sufficient for numerous crystallization attempts. It is unclear why these crystal trials failed to reproducibly yield crystals, however several suggestions could be made for future attempts. In principle, glycosylation sites may inhibit crystallization, and these (either on EPCR or LES γδ TCR) could be mutated. Alternatively, crystallization "rescue" approaches (e.g. lysine methylation) have previously been employed to crystallise intransigent proteins (Walter *et al*, 2006), and these could be attempted. Conceivably the low affinity of the LES/EPCR interaction may hamper TCR/ligand complex crystallization. Use of high affinity EPCR mutants (as identified in Chapter 4) in crystal trials with native sequence LES γδ TCR might circumvent this issue, however unfortunately to date such forms of EPCR have proven less soluble than wild-type EPCR. Finally, alterations in crystal packing could potentially aid crystallization. This could involve either addition of an antibody against one of the proteins, or engineering additional domains (e.g. onto EPCR) to facilitate crystal packing. Clearly these attempts could collectively represent a major undertaking, but potentially the level of information gained about the LES/EPCR interaction could be very substantial.

In principle, the use of recombinant γδ TCRs could also be extended beyond the uses mentioned above to ligand identification itself. By comparison, a range of cell surface receptors (e.g. CD48 family members, Podoplanin, other C-lectin receptors) have been used directly to help identify counter-ligands on the surface of target cells (Brown *et al*, 1998; Christou *et al*, 2008; Sancho *et al*, 2009) Conceivably, specific γδ TCRs could be produced in tetrameric/multimeric forms and stained against a panel of cell lines for which in-depth transcriptional analyses have been carried out. In principle, this could provide a direct read-out for ligand expression, which could then be correlated with transcriptional data to derive a

set of candidate ligands for each TCR, to be subsequently confirmed or excluded by direct binding analyses. While this strategy is likely to be attempted in collaboration with Professor Simon Davis, several indications highlight it is likely to be technically challenging. In addition to the fact it has never been carried out before, as an example, the LES γδ TCR has proven impossible to use successfully in tetrameric/multimeric form for staining of EPCR-expressing target cells. One potential reason for this could be the rather low affinity of the LES γδ TCR/EPCR interaction (see below). It is unclear whether this low affinity is unusual or not for human γδ TCR interactions. Moreover, the density of target γδ TCR antigen on target cells is likely to be variable, and low density ligands are likely to prove extremely challenging for TCR staining approaches, even though their expression might be sufficient for TCR activation, since T cells can be extremely sensitive to low levels of antigen. In my view, these attempts at direct recombinant TCR-based ligand identification are worthwhile but likely to be difficult, and alternative approaches such as blocking antibody generation (as for EPCR) (Willcox *et al*, 2012), and transcriptional approaches combined with activation assays of target cell recognition may be more productive routes forward.

In Chapter 4, I focussed on determining the binding mode of the LES γδ TCR/EPCR interaction, following on from parallel work in the Willcox laboratory by Dr Carrie Willcox that identified EPCR as a direct LES γδ TCR ligand (Willcox *et al*, 2012). One of the reasons for investigating this was the observation that EPCR expression on the target cell surface was necessary but not sufficient for productive LES T cell recognition. Since EPCR can bind lipids in an analogous fashion to CD1 molecules (Oganesyan *et al*, 2002), a possible explanation for this was that EPCR acted as an antigen presentation molecule for LES γδ T cells, with relevant "stress lipids" only presented in sufficient quantities on activatory cells.

Despite the fact that analogous changes in sphingolipid metabolism have been noted in both CMV infection and transformation (Machesky *et al*, 2008; Oskouian *et al*, 2007; Oskouian *et al*, 2006)the results in Chapter 4 defined LES/EPCR interaction as independent of the EPCR-bound lipid moiety. In addition they confirmed that, the interaction is independent of glycosylation but dependent on the conformation of the EPCR molecule (Willcox *et al*, 2012). These results suggest that rather than being recognized as an antigen presentation molecule, EPCR recognition by LES (and the activatory status of target cells) is dependent on the level of EPCR expression in combination with other factors extrinsic to EPCR. Of interest, the EPCR expression level on fibroblasts is unaffected by CMV infection, however experiments by Dr Déchanet-Merville's laboratory confirmed that CMV infection does upregulate certain adhesion/costimulatory molecules, including in particular ICAM-1, and that these stimulate LES T cell recognition. In the context of cellular transformation, while EPCR was again necessary for productive LES T cell recognition, the level of EPCR expression did not correlate well with activatory status. However, a range of publications have highlighted upregulation of EPCR on transformed epithelial cells (Scheffer *et al*, 2002; Tsuneyoshi *et al*, 2001; Anton *et al*, 2012) and unpublished data from the Willcox laboratory indicate upregulation in primary tumour tissue from a range of different cancers, suggesting increases in EPCR levels on tumours could be relevant to γδ T cell responses against tumours.

While the antigenic specificity of the LES T cell clone appears to be relatively rare (since it has not been observed in other patients, although it was expanded to high levels in the patient from which it derived), it is tempting to speculate what the findings outlined above mean for γδ T cell responses in general. Firstly, the observation that EPCR is an endothelial protein suggests Vδ2-negative T cells, which expand considerably in CMV-infected patients, might

use proteins such as EPCR to carry out immunosurveillance of the endothelium, which is a key site of CMV infection. For LES, it appears that the TCR ligand, EPCR, is constitutively present on endothelium, whereas altered expression of costimulatory molecules is crucial for stimulating γδ T cell effector responses. In this particular case, EPCR could therefore be viewed as just one part of a "multimolecular stress signature". This raises many questions, including in particular whether other Vδ2-negative γδ T cells utilise a similar mechanism (involving TCR recognition of self-antigens) to survey the endothelium for signs of viral infection. If so, LES γδ TCR could potentially provide a paradigm for many other γδ TCR reactivities used in CMV infection, even if the target antigens may be different, and possibly diverse. In keeping with with my observations and with this suggestion, are findings from a number of studies in mice, that indicate γδ T cells do not undergo stringent negative selection during thymic development (unlike αβ T cells), consistent with the possibility of self-antigen recognition in the periphery (Crowley *et al*, 2000) An alternative to this model of self-antigen recognition might be that during CMV infection, a portion of the γδ T cell repertoire could make responses to infected cells by directly recognizing viral proteins expressed at the surface, using their γδ TCR. While such responses would not help explain the observed dual reactivity of CMV-reactive Vδ2-negative T cells to tumour cells, consistent with this suggestion is a recent study by Chien and colleagues, who explored γδ T cell recognition of an algal protein (phycoerythrin, PE), concluding that specific CDR3 sequences in both the γ and δ chain can confer reactivity to PE, in the context of multiple different Vγ and Vδ regions (Zeng *et al*, 2012). Clearly many further studies are required to shed light on both of these potential models. However, irrespective of the molecular targets recognized by their TCRs, the potent effector functions of Vδ2-negative γδ T cells suggest they may play an important role in controlling CMV infection *in vivo*, particularly when αβ responses are compromised.

Defining the costimulatory ligands involved in stimulating γδ T cells in such contexts is also an area for additional study, although the experiments mentioned in Chapter 4 by Dr Déchanet's laboratory have already highlighted ICAM-1 as potentially important.

In addition to CMV infection, it is interesting to speculate on the specific and general relevance of EPCR recognition by LES in the context of anti-tumour immune responses. Previous work from Dr Déchanet-Merville's laboratory highlighted that renal transplant patients with CMV-induced expansions of the Vδ2-negative repertoire have significantly lower rates of post-transplant skin cancer, suggesting a possible role in tumour immunosurveillance (Couzi *et al*, 2010). In this context, the range of cellular targets recognized by γδ T cells, and their relationship to tumourigenesis, is of major interest. Conceivably EPCR might simply play a role for the immune system in this context, by analogy to NKG2D ligands to which it is homologous. However, its widespread upregulation in a number of tumours suggests its aberrant expression on tumour cells might be helpful to the tumour in some way. Of relevance is the normal physiological role of EPCR on endothelium, which is not only to suppress coagulation (by binding to Activated Protein C), but also to protect the endothelium from insults, such as p53-mediated apoptosis induced by hypoxia (Cheng *et al*, 2003). Given this, it is possible that aberrant EPCR expression on tumour cells enables them to evade apoptosis, and possibly to circumvent key stresses in the tumour microenvironment – such as pro-hypoxic stimuli. These ideas are the subject of ongoing work in the Willcox laboratory.

In Chapters 5 and 6, I attempted structural studies on Skint-1, the prototypic member of a multiprotein family that bears homology to both the butyrophilin and B7 families. Skint-1, the

gene for which was discovered following its serendipitous mutation in strains of Tacoma mice, was found to play a critical role in the selection of mouse Vγ5Vδ1 intraepithelial lymphocytes, typically referred to as Dendritic Epidermal T cells (DETC), whereas selection of other γδ T cell subsets was independent of Skint-1 (Boyden *et al*, 2008). In addition, since other Skint family members cannot substitute for Skint-1 expression in RTOC models of DETC maturation, it appears that the role of Skint-1 is non-redundant (Barbee *et al*, 2011). These data, together with its essential expression in the thymus on epithelium, and presence in the periphery on keratinocytes, establish Skint-1 is a candidate ligand for the Vγ5Vδ1 TCR. Moreover, recent studies suggest that the Vγ5Vδ1 TCR expressed on DETC may be involved in engagement of ligands constitutively expressed on the surface of keratinocytes prior to stress induction, consistent with the properties of Skint-1 (Chodaczek *et al*, 2012). The aim of Chapters 5 and 6 was to start to define the molecular features of Skint-1, with a view to understanding its functions at a molecular level. Although a structure determination of the two immunoglobulin-like ectodomains of Skint-1 proved impossible, experiments in Chapter 5 indicated that the membrane-distal Ig variable-like domain (IgV) could be produced in quantities sufficient for structural studies using X-ray crystallography and Nuclear Magnetic Resonance (NMR). Despite crystallizing the Skint-1 IgV domain, the crystals diffracted to low resolution (4Å), preventing structure determination using X-ray Crystallography. However, the structure of the membrane-distal domain of Skint-1 was successfully determined in solution using NMR.

While the solution structure of Skint-1 IgV inevitably only provides quite limited information, a few key conclusions can be drawn from this study. Firstly, the structure reveals several differences compared to homologues, a key comparison being with Myelin Oligodendrocyte

Glycoprotein (MOG), itself a homolog of the butyrophilins. This revealed two sizeable regions of high flexibility, as well as secondary structural differences, including a loop that is important in mediating dimerisation of MOG, but whose conformation in Skint-1 prevents dimerisation, consistent with size exclusion experiments that indicate the Skint-1 IgV domain is monomeric in solution. Secondly, comparisons with key family members provided two key insights. An initial point is that most of the amino acid differences between Skint-1, 2 and 7 are located at solvent-exposed residues. Consequently, the structure of Skint-1 IgV represents a first order mode of the IgV domain of both Skint-2 and Skint-7. However, an additional point is that despite a similar overall architecture for these domains, the nature of the solvent exposed changes suggested highly distinct electrostatic surfaces, with Skint-1 far less electropositive than these other two family members. A third key outcome of these structural studies, based on the fact that some residues identified from RTOC experiments as critical to Skint-1 function are entirely solvent exposed, is the identification of a putative binding site on the Skint-1 surface. Consistent with the non-redundant role of Skint-1 in mediating DETC maturation, the region involved contained several Skint-1-specific features, including a critical amino acid (D107), a distinctive hydrophobic patch, and a key proline residue. While not conclusive, these features are highly suggestive of a Skint-1-specific interaction with a counter-receptor.

While these data are interesting, and outline molecular features that may in time prove key to Skint-1's non-redundant function, they inevitably leave our picture of Skint-1 incomplete. This is emphasized partly by the results of RTOC experiments that highlight several other regions of the Skint-1 molecule (including the IgC-like domain, and the region just N-terminal of it) as critical to its function (Barbee *et al*, 2011). Undoubtedly lacking is a clear

understanding of the mechanism whereby Skint-1 mediates maturation of the DETC subset; a significant advance would be to know if Skint-1 is the ligand for the Vγ5Vδ1 TCR. Although this was beyond the scope of this thesis due to time constraints, these experiments are ongoing, using both BIAcore and Alphascreen approaches, and TCR expression methods similar to those described in Chapter 3. In addition, other questions, such as whether Skint-1 functions alone or in a multimolecular complex, also remain unresolved. DETC remain a highly intriguing example of tissue-resident γδ T cells, and their apparently constitutive recognition of TCR ligands on the surface of keratinocytes potentially represents a novel paradigm in γδ recognition of cellular stress. Clearly many further studies are required to outline the molecular details involved, and how these processes compare with human γδ T cell stress surveillance.

References

Adams, E.J., Chien, Y.H., and Garcia, K.C. (2005). Structure of a gammadelta T cell receptor in complex with the nonclassical MHC T22. Science *308*, 227-231.

Adams, E.J., Strop, P., Shin, S., Chien, Y.H., and Garcia, K.C. (2008). An autonomous CDR3delta is sufficient for recognition of the nonclassical MHC class I molecules T10 and T22 by gammadelta T cells. Nat immunol *9*, 777-784.

Allison, T.J., Winter, C.C., Fournie, J.J., Bonneville, M., and Garboczi, D.N. (2001). Structure of a human gammadelta T-cell antigen receptor. Nature *411*, 820-824.

Anton, I., Molina, E., Luis-Ravelo, D., Zandueta, C., Valencia, K., Ormazabal, C., Martinez-Canarias, S., Perurena, N., Pajares, M.J., Agorreta, J., *et al.* (2012). Receptor of activated protein C promotes metastasis and correlates with clinical outcome in lung adenocarcinoma. Am J Respir Crit Care Med *186*, 96-105.

Arden, B., Clark, S.P., Kabelitz, D., and Mak, T.W. (1995a). Human T-cell receptor variable gene segment families. Immunogenetics *42*, 455-500.

Arden, B., Clark, S.P., Kabelitz, D., and Mak, T.W. (1995b). Mouse T-cell receptor variable gene segment families. Immunogenetics *42*, 501-530.

Barbee, S.D., Woodward, M.J., Turchinovich, G., Mention, J.J., Lewis, J.M., Boyden, L.M., Lifton, R.P., Tigelaar, R., and Hayday, A.C. (2011). Skint-1 is a highly specific, unique selecting component for epidermal T cells. Proc Natl Acad Sci U S A *108*, 3330-3335.

Bauer, S., Groh, V., Wu, J., Steinle, A., Phillips, J.H., Lanier, L.L., and Spies, T. (1999). Activation of NK cells and T cells by NKG2D, a receptor for stress-inducible MICA. Science *285*, 727-729.

Bergfors, T. (2007). Screening and optimization methods for nonautomated crystallization laboratories. Methods in molecular biology *363*, 131-151.

Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., and Wiley, D.C. (1987). Structure of the human class I histocompatibility antigen, HLA-A2. Nature *329*, 506-512.

Blundell, T.L.J., L (1976). Protein Crystallography (Molecular Biology Series).

Boismenu, R., Feng, L., Xia, Y.Y., Chang, J.C., and Havran, W.L. (1996). Chemokine expression by intraepithelial gamma delta T cells. Implications for the recruitment of inflammatory cells to damaged epithelia. J Immunol *157*, 985-992.

Boismenu, R., and Havran, W.L. (1994). Modulation of epithelial cell growth by intraepithelial gamma delta T cells. Science *266*, 1253-1255.

Boniface, J.J., Reich, Z., Lyons, D.S., and Davis, M.M. (1999). Thermodynamics of T cell receptor binding to peptide-MHC: evidence for a general mechanism of molecular scanning. Proc Natl Acad Sci U S A *96*, 11446-11451.

Bonneville, M., Ishida, I., Itohara, S., Verbeek, S., Berns, A., Kanagawa, O., Haas, W., and Tonegawa, S. (1990). Self-tolerance to transgenic gamma delta T cells by intrathymic inactivation. Nature *344*, 163-165.

Boullier, S., Poquet, Y., Halary, F., Bonneville, M., Fournie, J.J., and Gougeon, M.L. (1998). Phosphoantigen activation induces surface translocation of intracellular CD94/NKG2A class I receptor on CD94- peripheral Vgamma9 Vdelta2 T cells but not on CD94- thymic or mature gammadelta T cell clones. Eur J Immunol *28*, 3399-3410.

Boulter, J.M., Schmitz, N., Sewell, A.K., Godkin, A.J., Bachmann, M.F., and Gallimore, A.M. (2007). Potent T cell agonism mediated by a very rapid TCR/pMHC interaction. Eur J Immunol *37*, 798-806.

Boyden, L.M., Lewis, J.M., Barbee, S.D., Bas, A., Girardi, M., Hayday, A.C., Tigelaar, R.E., and Lifton, R.P. (2008). Skint1, the prototype of a newly identified immunoglobulin superfamily gene cluster, positively selects epidermal gammadelta T cells. Nature genetics *40*, 656-662.

Breithaupt, C., Schafer, B., Pellkofer, H., Huber, R., Linington, C., and Jacob, U. (2008). Demyelinating myelin oligodendrocyte glycoprotein-specific autoantibody response is focused on one dominant conformational epitope region in rodents. J Immunol *181*, 1255-1263.

Brenner, M.B., McLean, J., Dialynas, D.P., Strominger, J.L., Smith, J.A., Owen, F.L., Seidman, J.G., Ip, S., Rosen, F., and Krangel, M.S. (1986). Identification of a putative second T-cell receptor. Nature *322*, 145-149.

Brigl, M., and Brenner, M.B. (2004). CD1: antigen presentation and T cell function. Annu Rev Immunol *22*, 817-890.

Brown, M.H., Boles, K., van der Merwe, P.A., Kumar, V., Mathew, P.A., and Barclay, A.N. (1998). 2B4, the natural killer and T cell immunoglobulin superfamily surface protein, is a ligand for CD48. J Exp Med *188*, 2083-2090.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., *et al.* (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta crystallographica *54*, 905-921.

Bukowski, J.F., Morita, C.T., Band, H., and Brenner, M.B. (1998). Crucial role of TCR gamma chain junctional region in prenyl pyrophosphate antigen recognition by gamma delta T cells. J Immunol *161*, 286-293.

Bukowski, J.F., Morita, C.T., Tanaka, Y., Bloom, B.R., Brenner, M.B., and Band, H. (1995). V gamma 2V delta 2 TCR-dependent recognition of non-peptide antigens and Daudi cells analyzed by TCR gene transfer. J Immunol *154*, 998-1006.

Carreno, L.J., Gonzalez, P.A., and Kalergis, A.M. (2006). Modulation of T cell function by TCR/pMHC binding kinetics. Immunobiology *211*, 47-64.

Cerwenka, A., Bakker, A.B., McClanahan, T., Wagner, J., Wu, J., Phillips, J.H., and Lanier, L.L. (2000). Retinoic acid early inducible genes define a ligand family for the activating NKG2D receptor in mice. Immunity *12*, 721-727.

Chang, V.T., Crispin, M., Aricescu, A.R., Harvey, D.J., Nettleship, J.E., Fennelly, J.A., Yu, C., Boles, K.S., Evans, E.J., Stuart, D.I., *et al.* (2007). Glycoprotein structural genomics: solving the glycosylation problem. Structure *15*, 267-273.

Chen, Y., Liu, P., Gao, F., Cheng, H., Qi, J., and Gao, G.F. (2010). A dimeric structure of PD-L1: functional units or evolutionary relics? Protein & cell *1*, 153-160.

Cheng, T., Liu, D., Griffin, J.H., Fernandez, J.A., Castellino, F., Rosen, E.D., Fukudome, K., and Zlokovic, B.V. (2003). Activated protein C blocks p53-mediated apoptosis in ischemic human brain endothelium and is neuroprotective. Nature medicine *9*, 338-342.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., and Thompson, J.D. (2003). Multiple sequence alignment with the Clustal series of programs. Nucleic acids research *31*, 3497-3500.

Chien, Y.H., Iwashima, M., Wettstein, D.A., Kaplan, K.B., Elliott, J.F., Born, W., and Davis, M.M. (1987). T-cell receptor delta gene rearrangements in early thymocytes. Nature *330*, 722-727.

Chodaczek, G., Papanna, V., Zal, M.A., and Zal, T. (2012). Body-barrier surveillance by epidermal gammadelta TCRs. Nature immunology *13*, 272-282.

Choi, S.P., Park, Y.C., Lee, J., Sim, S.J., and Chang, H.N. (2012). Effects of L-arginine on refolding of lysine-tagged human insulin-like growth factor 1 expressed in Escherichia coli. Bioprocess and biosystems engineering *35*, 255-263.

Choudhuri, K., and van der Merwe, P.A. (2007). Molecular mechanisms involved in T cell receptor triggering. Seminars in immunology *19*, 255-261.

Christou, C.M., Pearce, A.C., Watson, A.A., Mistry, A.R., Pollitt, A.Y., Fenton-May, A.E., Johnson, L.A., Jackson, D.G., Watson, S.P., and O'Callaghan, C.A. (2008). Renal cells activate the platelet receptor CLEC-2 through podoplanin. Biochem J *411*, 133-140.

Clark, S.P., Arden, B., Kabelitz, D., and Mak, T.W. (1995). Comparison of human and mouse T-cell receptor variable gene segment subfamilies. Immunogenetics *42*, 531-540.

Clements, C.S., Reid, H.H., Beddoe, T., Tynan, F.E., Perugini, M.A., Johns, T.G., Bernard, C.C., and Rossjohn, J. (2003). The crystal structure of myelin oligodendrocyte glycoprotein, a key autoantigen in multiple sclerosis. Proc Natl Acad Sci U S A *100*, 11059-11064.

Constant, P., Davodeau, F., Peyrat, M.A., Poquet, Y., Puzo, G., Bonneville, M., and Fournie, J.J. (1994). Stimulation of human gamma delta T cells by nonpeptidic mycobacterial ligands. Science *264*, 267-270.

Cornilescu, G., Delaglio, F., and Bax, A. (1999). Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR *13*, 289-302.

Couzi, L., Lafarge, X., Pitard, V., Neau-Cransac, M., Dromer, C., Billes, M.A., Lacaille, F., Moreau, J.F., Merville, P., and Dechanet-Merville, J. (2011). Gamma-delta T cell expansion is closely associated with cytomegalovirus infection in all solid organ transplant recipients. Transpl Int *24*, e40-42.

Couzi, L., Levaillant, Y., Jamai, A., Pitard, V., Lassalle, R., Martin, K., Garrigue, I., Hawchar, O., Siberchicot, F., Moore, N., *et al.* (2010). Cytomegalovirus-induced gammadelta T cells associate with reduced cancer risk after kidney transplantation. J Am Soc Nephrol *21*, 181-188.

Couzi, L., Pitard, V., Sicard, X., Garrigue, I., Hawchar, O., Merville, P., Moreau, J.F., and Dechanet-Merville, J. (2012). Antibody-dependent anti-cytomegalovirus activity of human gammadelta T cells expressing CD16 (FcgammaRIIIa). Blood *119*, 1418-1427.

Cresswell, P., Bangia, N., Dick, T., and Diedrich, G. (1999). The nature of the MHC class I peptide loading complex. Immunol Rev *172*, 21-28.

Crowley, M.P., Fahrer, A.M., Baumgarth, N., Hampl, J., Gutgemann, I., Teyton, L., and Chien, Y. (2000). A population of murine gammadelta T cells that recognize an inducible MHC class Ib molecule. Science *287*, 314-316.

D'Souza, C.D., Cooper, A.M., Frank, A.A., Mazzaccaro, R.J., Bloom, B.R., and Orme, I.M. (1997). An anti-inflammatory role for gamma delta T lymphocytes in acquired immunity to Mycobacterium tuberculosis. J Immunol *158*, 1217-1221.

Davies, D.R., and Metzger, H. (1983). Structural basis of antibody function. Annu Rev Immunol *1*, 87-117.

Davis, M.M., and Bjorkman, P.J. (1988). T-cell antigen receptor genes and T-cell recognition. Nature *334*, 395-402.

Davis, S.J., and van der Merwe, P.A. (2006). The kinetic-segregation model: TCR triggering and beyond. Nat immunol *7*, 803-809.

De Maria, A., Ferrazin, A., Ferrini, S., Ciccone, E., Terragna, A., and Moretta, L. (1992). Selective increase of a subset of T cell receptor gamma delta T lymphocytes in the peripheral blood of patients with human immunodeficiency virus type 1 infection. J Infect Dis *165*, 917-919.

Dechanet, J., Merville, P., Berge, F., Bone-Mane, G., Taupin, J.L., Michel, P., Joly, P., Bonneville, M., Potaux, L., and Moreau, J.F. (1999a). Major expansion of gammadelta T lymphocytes following cytomegalovirus infection in kidney allograft recipients. J Infect Dis *179*, 1-8.

Dechanet, J., Merville, P., Lim, A., Retiere, C., Pitard, V., Lafarge, X., Michelson, S., Meric, C., Hallet, M.M., Kourilsky, P., *et al.* (1999b). Implication of gammadelta T cells in the human immune response to cytomegalovirus. J Clin Invest *103*, 1437-1449.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR *6*, 277-293.

Dent, A.L., Matis, L.A., Hooshmand, F., Widacki, S.M., Bluestone, J.A., and Hedrick, S.M. (1990). Self-reactive gamma delta T cells are eliminated in the thymus. Nature *343*, 714-719.

Dessen, A., Lawrence, C.M., Cupo, S., Zaller, D.M., and Wiley, D.C. (1997). X-ray crystal structure of HLA-DR4 (DRA*0101, DRB1*0401) complexed with a peptide from human collagen II. Immunity *7*, 473-481.

Diefenbach, A., Jamieson, A.M., Liu, S.D., Shastri, N., and Raulet, D.H. (2000). Ligands for the murine NKG2D receptor: expression by tumor cells and activation of NK cells and macrophages. Nat Immunol *1*, 119-126.

Dieli, F., Gebbia, N., Poccia, F., Caccamo, N., Montesano, C., Fulfaro, F., Arcara, C., Valerio, M.R., Meraviglia, S., Di Sano, C., *et al.* (2003). Induction of gammadelta T-lymphocyte effector functions by bisphosphonate zoledronic acid in cancer patients in vivo. Blood *102*, 2310-2311.

Dieli, F., Vermijlen, D., Fulfaro, F., Caccamo, N., Meraviglia, S., Cicero, G., Roberts, A., Buccheri, S., D'Asaro, M., Gebbia, N., *et al.* (2007). Targeting human {gamma}delta} T cells with zoledronate and interleukin-2 for immunotherapy of hormone-refractory prostate cancer. Cancer research *67*, 7450-7457.

Ding, Y.H., Baker, B.M., Garboczi, D.N., Biddison, W.E., and Wiley, D.C. (1999). Four A6-TCR/peptide/HLA-A2 structures that generate very different T cell signals are nearly identical. Immunity *11*, 45-56.

Eberl, M., Hintz, M., Reichenberg, A., Kollas, A.K., Wiesner, J., and Jomaa, H. (2003). Microbial isoprenoid biosynthesis and human gammadelta T cell activation. FEBS letters *544*, 4-10.

Elliott, J.F., Rock, E.P., Patten, P.A., Davis, M.M., and Chien, Y.H. (1988). The adult T-cell receptor delta-chain is diverse and distinct from that of fetal thymocytes. Nature *331*, 627-631.

Ericsson, U.B., Hallberg, B.M., Detitta, G.T., Dekker, N., and Nordlund, P. (2006). Thermofluor-based high-throughput stability optimization of proteins for structural studies. Anal Biochem *357*, 289-298.

Evans, P., and McCoy, A. (2008). An introduction to molecular replacement. Acta crystallographica *64*, 1-10.

Gallagher, T., Alexander, P., Bryan, P., and Gilliland, G.L. (1994). Two crystal structures of the B1 immunoglobulin-binding domain of streptococcal protein G and comparison with NMR. Biochemistry *33*, 4721-4729.

Gao, G.F., Tormo, J., Gerth, U.C., Wyer, J.R., McMichael, A.J., Stuart, D.I., Bell, J.I., Jones, E.Y., and Jakobsen, B.K. (1997). Crystal structure of the complex between human CD8alpha(alpha) and HLA-A2. Nature *387*, 630-634.

Garboczi, D.N., Ghosh, P., Utz, U., Fan, Q.R., Biddison, W.E., and Wiley, D.C. (1996a). Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. Nature *384*, 134-141.

Garboczi, D.N., Hung, D.T., and Wiley, D.C. (1992). HLA-A2-peptide complexes: refolding and crystallization of molecules expressed in Escherichia coli and complexed with single antigenic peptides. Proc Nat Acad Sci U S A *89*, 3429-3433.

Garboczi, D.N., Utz, U., Ghosh, P., Seth, A., Kim, J., VanTienhoven, E.A., Biddison, W.E., and Wiley, D.C. (1996b). Assembly, specific binding, and crystallization of a human TCR-alphabeta with an antigenic Tax peptide from human T lymphotropic virus type 1 and the class I MHC molecule HLA-A2. J Immunol *157*, 5403-5410.

Garcia, K.C., Degano, M., Speir, J.A., and Wilson, I.A. (1999). Emerging principles for T cell receptor recognition of antigen in cellular immunity. Reviews in immunogenetics *1*, 75-90.

Garcia, K.C., Scott, C.A., Brunmark, A., Carbone, F.R., Peterson, P.A., Wilson, I.A., and Teyton, L. (1996). CD8 enhances formation of stable T-cell receptor/MHC class I molecule complexes. Nature *384*, 577-581.

Garcia, K.C., Tallquist, M.D., Pease, L.R., Brunmark, A., Scott, C.A., Degano, M., Stura, E.A., Peterson, P.A., Wilson, I.A., and Teyton, L. (1997). Alphabeta T cell receptor interactions with syngeneic and allogeneic ligands: affinity measurements and crystallization. Proc Nat Acad Sci U S A *94*, 13838-13843.

Garner, L.I., Salim, M., Mohammed, F., and Willcox, B.E. (2006). Expression, purification, and refolding of the myeloid inhibitory receptor leukocyte immunoglobulin-like receptor-5 for structural and ligand identification studies. Protein Expr Purif *47*, 490-497.

Girardi, M., Lewis, J., Glusac, E., Filler, R.B., Geng, L., Hayday, A.C., and Tigelaar, R.E. (2002). Resident skin-specific gammadelta T cells provide local, nonredundant regulation of cutaneous inflammation. J Exp Med *195*, 855-867.

Girardi, M., Oppenheim, D.E., Steele, C.R., Lewis, J.M., Glusac, E., Filler, R., Hobby, P., Sutton, B., Tigelaar, R.E., and Hayday, A.C. (2001). Regulation of cutaneous malignancy by gammadelta T cells. Science *294*, 605-609.

Goddard TD, K.D. (2004). SPARKY 3.

Groh, V., Bahram, S., Bauer, S., Herman, A., Beauchamp, M., and Spies, T. (1996). Cell stress-regulated human major histocompatibility complex class I gene expressed in gastrointestinal epithelium. Proc Nat Acad Sci U S A *93*, 12445-12450.

Groh, V., Rhinehart, R., Secrist, H., Bauer, S., Grabstein, K.H., and Spies, T. (1999). Broad tumor-associated expression and recognition by tumor-derived gamma delta T cells of MICA and MICB. Proc Nat Acad Sci U S A *96*, 6879-6884.

Groh, V., Steinle, A., Bauer, S., and Spies, T. (1998). Recognition of stress-induced MHC molecules by intestinal epithelial gammadelta T cells. Science *279*, 1737-1740.

Guerra, N., Tan, Y.X., Joncker, N.T., Choy, A., Gallardo, F., Xiong, N., Knoblaugh, S., Cado, D., Greenberg, N.M., and Raulet, D.H. (2008). NKG2D-deficient mice are defective in tumor surveillance in models of spontaneous malignancy. Immunity *28*, 571-580.

Guntert, P. (2004). Automated NMR structure calculation with CYANA. Methods Mol Biol *278*, 353-378.

Halary, F., Pitard, V., Dlubek, D., Krzysiek, R., de la Salle, H., Merville, P., Dromer, C., Emilie, D., Moreau, J.F., and Dechanet-Merville, J. (2005). Shared reactivity of V{delta}2(neg) {gamma}{delta} T cells against cytomegalovirus-infected cells and tumor intestinal epithelial cells. J Exp Med *201*, 1567-1578.

Hampl, J., Schild, H., Litzenberger, C., Baron, M., Crowley, M.P., and Chien, Y.H. (1999). The specificity of a weak gamma delta TCR interaction can be modulated by the glycosylation of the ligand. J Immunol *163*, 288-294.

Haregewoin, A., Soman, G., Hom, R.C., and Finberg, R.W. (1989). Human gamma delta+ T cells respond to mycobacterial heat-shock protein. Nature *340*, 309-312.

Havran, W.L., and Allison, J.P. (1990). Origin of Thy-1+ dendritic epidermal cells of adult mice from fetal thymic precursors. Nature *344*, 68-70.

Havran, W.L., Chien, Y.H., and Allison, J.P. (1991). Recognition of self antigens by skin-derived T cells with invariant gamma delta antigen receptors. Science *252*, 1430-1432.

Hayday, A.C. (2000). [gamma][delta] cells: a right time and a right place for a conserved third way of protection. Annu Rev Immunol *18*, 975-1026.

Hayday, A.C. (2009). Gammadelta T cells and the lymphoid stress-surveillance response. Immunity *31*, 184-196.

Hayday, A.C., Diamond, D.J., Tanigawa, G., Heilig, J.S., Folsom, V., Saito, H., and Tonegawa, S. (1985). Unusual organization and diversity of T-cell receptor alpha-chain genes. Nature *316*, 828-832.

Hedrick, S.M., Cohen, D.I., Nielsen, E.A., and Davis, M.M. (1984a). Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. Nature *308*, 149-153.

Hedrick, S.M., Nielsen, E.A., Kavaler, J., Cohen, D.I., and Davis, M.M. (1984b). Sequence relationships between putative T-cell receptor polypeptides and immunoglobulins. Nature *308*, 153-158.

Hendrickson, W.A. (1991). Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. Science *254*, 51-58.

Hewitt, R.E., Lissina, A., Green, A.E., Slay, E.S., Price, D.A., and Sewell, A.K. (2005). The bisphosphonate acute phase response: rapid and copious production of proinflammatory cytokines by peripheral blood gd T cells in response to aminobisphosphonates is inhibited by statins. Clin Exp Immunol *139*, 101-111.

Hozumi, N., and Tonegawa, S. (1976). Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. Proc Nat Acad Sci U S A *73*, 3628-3632.

Huber, H., Descossy, P., van Brandwijk, R., and Knop, J. (1995). Activation of murine epidermal TCR-gamma delta+ T cells by keratinocytes treated with contact sensitizers. J Immunol *155*, 2888-2894.

Ito, K., Van Kaer, L., Bonneville, M., Hsu, S., Murphy, D.B., and Tonegawa, S. (1990). Recognition of the product of a novel MHC TL region gene (27b) by a mouse gamma delta T cell receptor. Cell *62*, 549-561.

Jagow, H.S.a.G.v. (1987). Tricine-Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis for the Separation of Proteins in the Range from 1 to 100 kDa. ANALYTICAL BIOCHEMISTRY *166*, 368-379.

Jameson, J., Ugarte, K., Chen, N., Yachi, P., Fuchs, E., Boismenu, R., and Havran, W.L. (2002). A role for skin gammadelta T cells in wound repair. Science *296*, 747-749.

Janeway, C.A., Jr., Jones, B., and Hayday, A. (1988). Specificity and function of T cells bearing gamma delta receptors. Immunol Today *9*, 73-76.

Janeway, C.A., Jr. Travers, P., Walport, M., Shlomchik, M. (2005). Immunobiology : the immune system in health and disease. *volume 5*.

Jensen, K.D., Su, X., Shin, S., Li, L., Youssef, S., Yamasaki, S., Steinman, L., Saito, T., Locksley, R.M., Davis, M.M., *et al.* (2008). Thymic selection determines gammadelta T cell effector fate: antigen-naive cells make interleukin-17 and antigen-experienced cells make interferon gamma. Immunity *29*, 90-100.

Johnson, R.M., Lancki, D.W., Sperling, A.I., Dick, R.F., Spear, P.G., Fitch, F.W., and Bluestone, J.A. (1992). A murine CD4-, CD8- T cell receptor-gamma delta T lymphocyte clone specific for herpes simplex virus glycoprotein I. J Immunol *148*, 983-988.

Kabsch, W. (2010). Integration, scaling, space-group assignment and post-refinement. Acta crystallographica *66*, 133-144.

Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers *22*, 2577-2637.

Kantardjieff, K.A., and Rupp, B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. Protein Sci *12*, 1865-1871.

Kapsenberg, M.L. (2009). Gammadelta T cell receptors without a job. Immunity *31*, 181-183.

Keir, M.E., Butte, M.J., Freeman, G.J., and Sharpe, A.H. (2008). PD-1 and its ligands in tolerance and immunity. Annu Rev Immunol *26*, 677-704.

King, D.P., Hyde, D.M., Jackson, K.A., Novosad, D.M., Ellis, T.N., Putney, L., Stovall, M.Y., Van Winkle, L.S., Beaman, B.L., and Ferrick, D.A. (1999). Cutting edge: protective response to pulmonary injury requires gamma delta T lymphocytes. J Immunol *162*, 5033-5036.

Kjer-Nielsen, L., Patel, O., Corbett, A.J., Le Nours, J., Meehan, B., Liu, L., Bhati, M., Chen, Z., Kostenko, L., Reantragoon, R., *et al.* (2012). MR1 presents microbial vitamin B metabolites to MAIT cells. Nature *491*, 717-723.

Knight, A., Madrigal, A.J., Grace, S., Sivakumaran, J., Kottaridis, P., Mackinnon, S., Travers, P.J., and Lowdell, M.W. (2010). The role of Vdelta2-negative gammadelta T cells during cytomegalovirus reactivation in recipients of allogeneic stem cell transplantation. Blood *116*, 2164-2172.

Koch, M., Stronge, V.S., Shepherd, D., Gadola, S.D., Mathew, B., Ritter, G., Fersht, A.R., Besra, G.S., Schmidt, R.R., Jones, E.Y., and Cerundolo, V. (2005). The crystal structure of human CD1d with and without alpha-galactosylceramide. Nature immunology *6*, 819-826.

Koradi, R., Billeter, M., and Wuthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graph *14*, 51-55, 29-32.

Korn, T., Bettelli, E., Oukka, M., and Kuchroo, V.K. (2009). IL-17 and Th17 Cells. Annu Rev Immunol *27*, 485-517.

Lafarge, X., Merville, P., Cazin, M.C., Berge, F., Potaux, L., Moreau, J.F., and Dechanet-Merville, J. (2001). Cytomegalovirus infection in transplant recipients resolves when circulating gammadelta T lymphocytes expand, suggesting a protective antiviral role. J Infect Dis *184*, 533-541.

Lafarge, X., Pitard, V., Ravet, S., Roumanes, D., Halary, F., Dromer, C., Vivier, E., Paul, P., Moreau, J.F., and Dechanet-Merville, J. (2005). Expression of MHC class I receptors confers functional intraclonal heterogeneity to a reactive expansion of gammadelta T cells. Eur J Immunol *35*, 1896-1905.

Laskowski RA, M.M., Moss DS, Thornton JM. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr *26*, 283–291.

Lesk, A.M., and Chothia, C. (1982). Evolution of proteins formed by beta-sheets. II. The core of the immunoglobulin domains. J Mol Biol *160*, 325-342.

Lewis, J.M., Girardi, M., Roberts, S.J., Barbee, S.D., Hayday, A.C., and Tigelaar, R.E. (2006). Selection of the cutaneous intraepithelial gammadelta+ T cell repertoire by a thymic stromal determinant. Nat Immunol *7*, 843-850.

Li, L., and Ismagilov, R.F. (2010). Protein crystallization using microfluidic technologies based on valves, droplets, and SlipChip. Annu Rev Biophys *39*, 139-158.

Li, P., Willie, S.T., Bauer, S., Morris, D.L., Spies, T., and Strong, R.K. (1999). Crystal structure of the MHC class I homolog MIC-A, a gammadelta T cell ligand. Immunity *10*, 577-584.

Lin, D.Y., Tanaka, Y., Iwasaki, M., Gittis, A.G., Su, H.P., Mikami, B., Okazaki, T., Honjo, T., Minato, N., and Garboczi, D.N. (2008). The PD-1/PD-L1 complex resembles the antigen-binding Fv domains of antibodies and T cell receptors. Proc Nat Acad Sci U S A *105*, 3011-3016.

Linder, M.E., and Deschenes, R.J. (2007). Palmitoylation: policing protein stability and traffic. Nature reviews *8*, 74-84.

Linge, J.P., O'Donoghue, S.I., and Nilges, M. (2001). Automated assignment of ambiguous nuclear overhauser effects with ARIA. Methods in enzymology *339*, 71-90.

Lockhart, E., Green, A.M., and Flynn, J.L. (2006). IL-17 production is dominated by gammadelta T cells rather than CD4 T cells during Mycobacterium tuberculosis infection. J Immunol *177*, 4662-4669.

Ma, Y., Aymeric, L., Locher, C., Mattarollo, S.R., Delahaye, N.F., Pereira, P., Boucontet, L., Apetoh, L., Ghiringhelli, F., Casares, N., *et al.* (2011). Contribution of IL-17-producing gamma delta T cells to the efficacy of anticancer chemotherapy. J Exp Med *208*, 491-503.

Machesky, N.J., Zhang, G., Raghavan, B., Zimmerman, P., Kelly, S.L., Merrill, A.H., Jr., Waldman, W.J., Van Brocklyn, J.R., and Trgovcich, J. (2008). Human cytomegalovirus regulates bioactive sphingolipids. J Biol Chem *283*, 26148-26160.

Madden, D.R., Garboczi, D.N., and Wiley, D.C. (1993). The antigenic identity of peptide-MHC complexes: a comparison of the conformations of five viral peptides presented by HLA-A2. Cell *75*, 693-708.

Mallick-Wood, C.A., Lewis, J.M., Richie, L.I., Owen, M.J., Tigelaar, R.E., and Hayday, A.C. (1998). Conservation of T cell receptor conformation in epidermal gammadelta cells with disrupted primary Vgamma gene usage. Science *279*, 1729-1733.

Matis, L.A., Fry, A.M., Cron, R.Q., Cotterman, M.M., Dick, R.F., and Bluestone, J.A. (1989). Structure and specificity of a class II MHC alloreactive gamma delta T cell receptor heterodimer. Science *245*, 746-749.

Matsuda, S., Kudoh, S., and Katayama, S. (2001). Enhanced formation of azoxymethane-induced colorectal adenocarcinoma in gammadelta T lymphocyte-deficient mice. Jpn J Cancer Res *92*, 880-885.

Matsui, K., Boniface, J.J., Reay, P.A., Schild, H., Fazekas de St Groth, B., and Davis, M.M. (1991). Low affinity interaction of peptide-MHC complexes with T cell receptors. Science *254*, 1788-1791.

Medzhitov, R., Preston-Hurlburt, P., and Janeway, C.A., Jr. (1997). A human homologue of the Drosophila Toll protein signals activation of adaptive immunity. Nature *388*, 394-397.

Meraviglia, S., Eberl, M., Vermijlen, D., Todaro, M., Buccheri, S., Cicero, G., La Mendola, C., Guggino, G., D'Asaro, M., Orlando, V., *et al.* (2010). In vivo manipulation of Vgamma9Vdelta2 T cells with zoledronate and low-dose interleukin-2 for immunotherapy of advanced breast cancer patients. Clin Exp Immunol *161*, 290-297.

Mistry, A.R., and O'Callaghan, C.A. (2007). Regulation of ligands for the activating receptor NKG2D. Immunology *121*, 439-447.

Momany, A.W.a.C. (2007). Addressing the Protein Crystallisation bottleneck by Cocrystallisation. Crystal Growth and Design *7*, 2219-2225.

Moore, T.A., Moore, B.B., Newstead, M.W., and Standiford, T.J. (2000). Gamma delta-T cells are critical for survival and early proinflammatory cytokine gene expression during murine Klebsiella pneumonia. J Immunol *165*, 2643-2650.

Morita, C.T., Beckman, E.M., Bukowski, J.F., Tanaka, Y., Band, H., Bloom, B.R., Golan, D.E., and Brenner, M.B. (1995). Direct presentation of nonpeptide prenyl pyrophosphate antigens to human gamma delta T cells. Immunity *3*, 495-507.

Morita, C.T., Lee, H.K., Wang, H., Li, H., Mariuzza, R.A., and Tanaka, Y. (2001). Structural features of nonpeptide prenyl pyrophosphates that determine their antigenicity for human gamma delta T cells. J Immunol *167*, 36-41.

Moser, B., and Brandes, M. (2006). Gammadelta T cells: an alternative type of professional APC. Trends in immunology *27*, 112-118.

Mucida, D., and Cheroutre, H. (2010). The many face-lifts of CD4 T helper cells. Adv Immunol *107*, 139-152.

Mukasa, A., Born, W.K., and O'Brien, R.L. (1999). Inflammation alone evokes the response of a TCR-invariant mouse gamma delta T cell subset. J Immunol *162*, 4910-4913.

Nakanishi, N., Maeda, K., Ito, K., Heller, M., and Tonegawa, S. (1987). T gamma protein is expressed on murine fetal thymocytes as a disulphide-linked heterodimer. Nature *325*, 720-723.

Nedellec, S., Sabourin, C., Bonneville, M., and Scotet, E. (2010). NKG2D costimulates human V gamma 9V delta 2 T cell antitumor cytotoxicity through protein kinase C theta-dependent modulation of early TCR-induced calcium and transduction signals. J Immunol *185*, 55-63.

Nicholls, S., Piper, K.P., Mohammed, F., Dafforn, T.R., Tenzer, S., Salim, M., Mahendra, P., Craddock, C., van Endert, P., Schild, H., *et al.* (2009). Secondary anchor polymorphism in the HA-1 minor histocompatibility antigen critically affects MHC stability and TCR recognition. Proc Nat Acad Sci U S A *106*, 3889-3894.

O'Callaghan C, A., Byford, M.F., Wyer, J.R., Willcox, B.E., Jakobsen, B.K., McMichael, A.J., and Bell, J.I. (1999). BirA enzyme: production and application in the study of membrane receptor-ligand interactions by site-specific biotinylation. Anal Biochem *266*, 9-15.

O'Callaghan, C.A., Cerwenka, A., Willcox, B.E., Lanier, L.L., and Bjorkman, P.J. (2001). Molecular competition for NKG2D: H60 and RAE1 compete unequally for NKG2D with dominance of H60. Immunity *15*, 201-211.

Oganesyan V, O.N., Terzyan S, Qu D, Dauter Z, Esmon NL, Esmon CT. (2002). The crystal structure of the endothelial protein C receptor and a bound phospholipid. *277(28)*, 24851-24854.

Oskouian, B., and Saba, J. (2007). Sphingosine-1-phosphate metabolism and intestinal tumorigenesis: lipid signaling strikes again. Cell cycle *6*, 522-527.

Oskouian, B., Sooriyakumaran, P., Borowsky, A.D., Crans, A., Dillard-Telm, L., Tam, Y.Y., Bandhuvula, P., and Saba, J.D. (2006). Sphingosine-1-phosphate lyase potentiates apoptosis via p53- and p38-dependent pathways and is down-regulated in colon cancer. Proc Nat Acad Sci U S A *103*, 17384-17389.

Ouellette, T., Destrau, S., Ouellette, T., Zhu, J., Roach, J.M., Coffman, J.D., Hecht, T., Lynch, J.E., and Giardina, S.L. (2003). Production and purification of refolded recombinant human IL-7 from inclusion bodies. Protein expression and purification *30*, 156-166.

Pamer, E., and Cresswell, P. (1998). Mechanisms of MHC class I--restricted antigen processing. Annu Rev immunol *16*, 323-358.

Pardoll, D.M., Fowlkes, B.J., Bluestone, J.A., Kruisbeek, A., Maloy, W.L., Coligan, J.E., and Schwartz, R.H. (1987). Differential expression of two distinct T-cell receptors during thymocyte development. Nature *326*, 79-81.

Peng, S.L., Madaio, M.P., Hayday, A.C., and Craft, J. (1996). Propagation and regulation of systemic autoimmunity by gammadelta T cells. J Immunol *157*, 5689-5698.

Pfeffer, K., Schoel, B., Gulle, H., Kaufmann, S.H., and Wagner, H. (1990). Primary responses of human T cells to mycobacteria: a frequent set of gamma/delta T cells are stimulated by protease-resistant ligands. Eur J Immunol *20*, 1175-1179.

Pitard, V., Roumanes, D., Lafarge, X., Couzi, L., Garrigue, I., Lafon, M.E., Merville, P., Moreau, J.F., and Dechanet-Merville, J. (2008). Long-term expansion of effector/memory Vdelta2-gammadelta T cells is a specific blood signature of CMV infection. Blood *112*, 1317-1324.

Pont, F., Familiades, J., Dejean, S., Fruchon, S., Cendron, D., Poupot, M., Poupot, R., L'Faqihi-Olive, F., Prade, N., Ycart, B., and Fournie, J.J. (2012). The gene expression profile of phosphoantigen-specific human gammadelta T lymphocytes is a blend of alphabeta T-cell and NK-cell signatures. Eur J Immunol *42*, 228-240.

Ramsburg, E., Tigelaar, R., Craft, J., and Hayday, A. (2003). Age-dependent requirement for gammadelta T cells in the primary but not secondary protective immune response against an intestinal parasite. J Exp Med *198*, 1403-1414.

Reinherz, E.L., Tan, K., Tang, L., Kern, P., Liu, J., Xiong, Y., Hussey, R.E., Smolyar, A., Hare, B., Zhang, R., *et al.* (1999). The crystal structure of a T cell receptor in complex with peptide and MHC class II. Science *286*, 1913-1921.

Roark, C.L., Simonian, P.L., Fontenot, A.P., Born, W.K., and O'Brien, R.L. (2008). gammadelta T cells: an important source of IL-17. Curr Opin Immunol *20*, 353-357.

Roberts, S.J., Smith, A.L., West, A.B., Wen, L., Findly, R.C., Owen, M.J., and Hayday, A.C. (1996). T-cell alpha beta + and gamma delta + deficient mice display abnormal but distinct phenotypes toward a natural, widespread infection of the intestinal epithelium. Proc Nat Acad Sci U S A *93*, 11774-11779.

Rocchia, W., Sridharan, S., Nicholls, A., Alexov, E., Chiabrera, A., and Honig, B. (2002). Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: applications to the molecular systems and geometric objects. J Comput Chem *23*, 128-137.

Rock, E.P., Sibbald, P.R., Davis, M.M., and Chien, Y.H. (1994). CDR3 length in antigen-specific immune receptors. J Exp Med *179*, 323-328.

Rudolph, M.G., Stanfield, R.L., and Wilson, I.A. (2006). How TCRs bind MHCs, peptides, and coreceptors. Annu Rev Immunol *24*, 419-466.

Rudolph, M.G., and Wilson, I.A. (2002). The specificity of TCR/pMHC interaction. Curr Opin Immunol *14*, 52-65.

Rudolph, M.G., Wingren, C., Crowley, M.P., Chien, Y.H., and Wilson, I.A. (2004). Combined pseudo-merohedral twinning, non-crystallographic symmetry and pseudo-translation in a monoclinic crystal form of the gammadelta T-cell ligand T10. Acta crystallographica *60*, 656-664.

Saito, H., Kranz, D.M., Takagaki, Y., Hayday, A.C., Eisen, H.N., and Tonegawa, S. (1984). A third rearranged and expressed gene in a clone of cytotoxic T lymphocytes. Nature *312*, 36-40.

Sancho, D., Joffre, O.P., Keller, A.M., Rogers, N.C., Martinez, D., Hernanz-Falcon, P., Rosewell, I., and Reis e Sousa, C. (2009). Identification of a dendritic cell receptor that couples sensing of necrosis to immunity. Nature *458*, 899-903.

Scheffer, G.L., Flens, M.J., Hageman, S., Izquierdo, M.A., Shoemaker, R.H., and Scheper, R.J. (2002). Expression of the vascular endothelial cell protein C receptor in epithelial tumour cells. Eur J Cancer *38*, 1535-1542.

Schild, H., and Chien, Y.H. (1994). The recognition of MHC molecules by gamma delta T cells. Behring Institute Mitteilungen, 113-123.

Schneiders, F.L., de Bruin, R.C., Santegoets, S.J., Bonneville, M., Scotet, E., Scheper, R.J., Verheul, H.M., de Gruijl, T.D., and van der Vliet, H.J. (2012). Activated iNKT cells promote Vgamma9Vdelta2-T cell anti-tumor effector functions through the production of TNF-alpha. Clinl Immunol *142*, 194-200.

Sciammas, R., Johnson, R.M., Sperling, A.I., Brady, W., Linsley, P.S., Spear, P.G., Fitch, F.W., and Bluestone, J.A. (1994). Unique antigen recognition by a herpesvirus-specific TCR-gamma delta cell. J Immunol *152*, 5392-5397.

Sherwood, A.M., Desmarais, C., Livingston, R.J., Andriesen, J., Haussler, M., Carlson, C.S., and Robins, H. (2011). Deep sequencing of the human TCRgamma and TCRbeta repertoires suggests that TCRbeta rearranges after alphabeta and gammadelta T cell commitment. Sc Transl Med *3*, 90ra61.

Shi, Y., Mowery, R.A., Ashley, J., Hentz, M., Ramirez, A.J., Bilgicer, B., Slunt-Brown, H., Borchelt, D.R., and Shaw, B.F. (2012). Abnormal SDS-PAGE migration of cytosolic proteins can identify domains and mechanisms that control surfactant binding. Protein Sci *21*, 1197-1209.

Shires, J., Theodoridis, E., and Hayday, A.C. (2001). Biological insights into TCRgammadelta+ and TCRalphabeta+ intraepithelial lymphocytes provided by serial analysis of gene expression (SAGE). Immunity *15*, 419-434.

Simpson, A.A., Mohammed, F., Salim, M., Tranter, A., Rickinson, A.B., Stauss, H.J., Moss, P.A., Steven, N.M., and Willcox, B.E. (2011). Structural and energetic evidence for highly peptide-specific tumor antigen targeting via allo-MHC restriction. Proc Nat Acad Sci U S A *108*, 21176-21181.

Sorensen, H.P. (2010). Towards universal systems for recombinant gene expression. Microbial cell factories *9*, 27.

Spada, F.M., Grant, E.P., Peters, P.J., Sugita, M., Melian, A., Leslie, D.S., Lee, H.K., van Donselaar, E., Hanson, D.A., Krensky, A.M., *et al.* (2000). Self-recognition of CD1 by gamma/delta T cells: implications for innate immunity. J Exp Med *191*, 937-948.

Srivastava, P. (2002). Interaction of heat shock proteins with peptides and antigen presenting cells: chaperoning of the innate and adaptive immune responses. Annu Rev Immunol *20*, 395-425.

Steele, C.R., Van Remoortere, K.C., and Hayday, A.C. (2003). Production of a soluble gammadelta T-cell receptor to identify ligands for the murine intestinal intraepithelial gammadelta T cell population. Journal of chromatography *786*, 297-304.

Strid, J., Roberts, S.J., Filler, R.B., Lewis, J.M., Kwong, B.Y., Schpero, W., Kaplan, D.H., Hayday, A.C., and Girardi, M. (2008). Acute upregulation of an NKG2D ligand promotes rapid reorganization of a local immune compartment with pleiotropic effects on carcinogenesis. Nature Immunol *9*, 146-154.

Stritesky, G.L., Jameson, S.C., and Hogquist, K.A. (2012). Selection of self-reactive T cells in the thymus. Annu Rev Immunol *30*, 95-114.

Susannah D. Barbeea, b., 1, Martin J. Woodwarda,b,1, Gleb Turchinovicha,b, Jean-Jacques Mentionb, Julia M. Lewisc,, and Lynn M. Boydend, e., Richard P. Liftond,e, Robert Tigelaarc, and Adrian C. Haydaya,b,2 (2011). Skint-1 is a highly specific, unique selecting component for epidermal T cells. PNAS *108*, 3330-3335.

Sutton, C.E., Lalor, S.J., Sweeney, C.M., Brereton, C.F., Lavelle, E.C., and Mills, K.H. (2009). Interleukin-1 and IL-23 induce innate IL-17 production from gammadelta T cells, amplifying Th17 responses and autoimmunity. Immunity *31*, 331-341.

Tanaka, Y., Sano, S., Nieves, E., De Libero, G., Rosa, D., Modlin, R.L., Brenner, M.B., Bloom, B.R., and Morita, C.T. (1994). Nonpeptide ligands for human gamma delta T cells. Proc Nat Acad Sci U S A *91*, 8175-8179.

Teitell, M., Cheroutre, H., Panwala, C., Holcombe, H., Eghtesady, P., and Kronenberg, M. (1994). Structure and function of H-2 T (Tla) region class I MHC molecules. Crit Rev Immunol *14*, 1-27.

Thoma-Uszynski, S., Ladel, C.H., and Kaufmann, S.H. (1997). Abscess formation in Listeria monocytogenes-infected gamma delta T cell deficient mouse mutants involves alpha beta T cells. Microb Patho *22*, 123-128.

Thornton, B.L.S.T.M. (1985). Beta-Hairpin families in globular proteins. Nature *136*.

Tilloy, F., Treiner, E., Park, S.H., Garcia, C., Lemonnier, F., de la Salle, H., Bendelac, A., Bonneville, M., and Lantz, O. (1999). An invariant T cell receptor alpha chain defines a novel TAP-independent major histocompatibility complex class Ib-restricted alpha/beta T cell subpopulation in mammals. J Exp Med *189*, 1907-1921.

Towbin, H., Staehelin, T., and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Nat Acad Sci U S A *76*, 4350-4354.

Townsend, A., and Bodmer, H. (1989). Antigen recognition by class I-restricted T lymphocytes. Annu Rev Immunol *7*, 601-624.

Townsend, A., and McMichael, A. (1987). MHC protein structure. Those images that yet fresh images beget. Nature *329*, 482-483.

Treiner, E., Duban, L., Bahram, S., Radosavljevic, M., Wanner, V., Tilloy, F., Affaticati, P., Gilfillan, S., and Lantz, O. (2003). Selection of evolutionarily conserved mucosal-associated invariant T cells by MR1. Nature *422*, 164-169.

Tsuneyoshi, N., Fukudome, K., Horiguchi, S., Ye, X., Matsuzaki, M., Toi, M., Suzuki, K., and Kimoto, M. (2001). Expression and anticoagulant function of the endothelial cell protein C receptor (EPCR) in cancer cell lines. Thromb Haemost *85*, 356-361.

Turchinovich, G., and Hayday, A.C. (2011). Skint-1 identifies a common molecular mechanism for the development of interferon-gamma-secreting versus interleukin-17-secreting gammadelta T cells. Immunity *35*, 59-68.

Umemura, M., Yahagi, A., Hamada, S., Begum, M.D., Watanabe, H., Kawakami, K., Suda, T., Sudo, K., Nakae, S., Iwakura, Y., and Matsuzaki, G. (2007). IL-17-mediated regulation of innate and acquired immune response against pulmonary Mycobacterium bovis bacille Calmette-Guerin infection. J Immunol *178*, 3786-3796.

van der Merwe, P.A., and Davis, S.J. (2003). Molecular interactions mediating T cell antigen recognition. Annu Rev Immunol *21*, 659-684.

Vermijlen, D., Brouwer, M., Donner, C., Liesnard, C., Tackoen, M., Van Rysselberge, M., Twite, N., Goldman, M., Marchant, A., and Willems, F. (2010). Human cytomegalovirus elicits fetal gammadelta T cell responses in utero. J Exp Med *207*, 807-821.

Vernooij, B.T., Lenstra, J.A., Wang, K., and Hood, L. (1993). Organization of the murine T-cell receptor gamma locus. Genomics *17*, 566-574.

Vignali, D.A., Collison, L.W., and Workman, C.J. (2008). How regulatory T cells work. Nat Rev Immunol *8*, 523-532.

Walter, T.S., Meier, C., Assenberg, R., Au, K.F., Ren, J., Verma, A., Nettleship, J.E., Owens, R.J., Stuart, D.I., and Grimes, J.M. (2006). Lysine methylation as a routine rescue strategy for protein crystallization. Structure *14*, 1617-1622.

Wang, H., Fang, Z., and Morita, C.T. (2010). Vgamma2Vdelta2 T Cell Receptor recognition of prenyl pyrophosphates is dependent on all CDRs. J Immunol *184*, 6209-6222.

Watts, C. (1997). Capture and processing of exogenous antigens for presentation on MHC molecules. Annu Rev Immunol *15*, 821-850.

Wells, F.B., Gahm, S.J., Hedrick, S.M., Bluestone, J.A., Dent, A., and Matis, L.A. (1991). Requirement for positive selection of gamma delta receptor-bearing T cells. Science *253*, 903-905.

Wen, L., Barber, D.F., Pao, W., Wong, F.S., Owen, M.J., and Hayday, A. (1998). Primary gamma delta cell clones can be defined phenotypically and functionally as Th1/Th2 cells and illustrate the association of CD4 with Th2 differentiation. J Immunol *160*, 1965-1974.

Wen, L., and Hayday, A.C. (1997). Gamma delta T-cell help in responses to pathogens and in the development of systemic autoimmunity. Immunol Res *16*, 229-241.

Wilhelm, M., Kunzmann, V., Eckstein, S., Reimer, P., Weissinger, F., Ruediger, T., and Tony, H.P. (2003). Gammadelta T cells for immune therapy of patients with lymphoid malignancies. Blood *102*, 200-206.

Willcox, B.E., Gao, G.F., Wyer, J.R., Ladbury, J.E., Bell, J.I., Jakobsen, B.K., and van der Merwe, P.A. (1999a). TCR binding to peptide-MHC stabilizes a flexible recognition interface. Immunity *10*, 357-365.

Willcox, B.E., Gao, G.F., Wyer, J.R., O'Callaghan, C.A., Boulter, J.M., Jones, E.Y., van der Merwe, P.A., Bell, J.I., and Jakobsen, B.K. (1999b). Production of soluble alphabeta T-cell receptor heterodimers suitable for biophysical analysis of ligand binding. Protein Sci *8*, 2418-2423.

Willcox, B.E., Thomas, L.M., Chapman, T.L., Heikema, A.P., West, A.P., Jr., and Bjorkman, P.J. (2002). Crystal structure of LIR-2 (ILT4) at 1.8 A: differences from LIR-1 (ILT2) in regions implicated in the binding of the Human Cytomegalovirus class I MHC homolog UL18. BMC structural biology *2*, 6.

Willcox, B.E., Willcox, C.R., Dover, L.G., and Besra, G. (2007). Structures and functions of microbial lipid antigens presented by CD1. Curr Top Microbiol Immunol *314*, 73-110.

Willcox, C.R., Pitard, V., Netzer, S., Couzi, L., Salim, M., Silberzahn, T., Moreau, J.F., Hayday, A.C., Willcox, B.E., and Dechanet-Merville, J. (2012). Cytomegalovirus and tumor stress surveillance by binding of a human gammadelta T cell antigen receptor to endothelial protein C receptor. Nature immunology *13*, 872-879.

Williams, A.F., and Barclay, A.N. (1988). The immunoglobulin superfamily--domains for cell surface recognition. Annu Rev Immunol *6*, 381-405.

Wilson, I.A., Stanfield, R.L., Rini, J.M., Arevalo, J.H., Schulze-Gahmen, U., Fremont, D.H., and Stura, E.A. (1991). Structural aspects of antibodies and antibody-antigen complexes. Ciba Found Symp *159*, 13-28; discussion 28-39.

Wingren, C., Crowley, M.P., Degano, M., Chien, Y., and Wilson, I.A. (2000). Crystal structure of a gammadelta T cell receptor ligand T22: a truncated MHC-like fold. Science *287*, 310-314.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., *et al.* (2011). Overview of the CCP4 suite and current developments. Acta crystallographica *67*, 235-242.

Wishart, D.S., and Sykes, B.D. (1994). Chemical shifts as a tool for structure determination. Methods in enzymology *239*, 363-392.

Wu, J., Groh, V., and Spies, T. (2002). T cell antigen receptor engagement and specificity in the recognition of stress-inducible MHC class I-related chains by human epithelial gamma delta T cells. J Immunol *169*, 1236-1240.

Wu, T.T., and Kabat, E.A. (1970). An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. J Exp Med *132*, 211-250.

Wucherpfennig, K.W. (2010). The first structures of T cell receptors bound to peptide-MHC. J Immunol *185*, 6391-6393.

Xu, B., Pizarro, J.C., Holmes, M.A., McBeth, C., Groh, V., Spies, T., and Strong, R.K. (2011). Crystal structure of a gammadelta T-cell receptor specific for the human MHC class I homolog MICA. Proc Nat Acad Sci U S A *108*, 2414-2419.

Yamasaki, S., Ishikawa, E., Sakuma, M., Ogata, K., Sakata-Sogawa, K., Hiroshima, M., Wiest, D.L., Tokunaga, M., and Saito, T. (2006). Mechanistic basis of pre-T cell receptor-mediated autonomous signaling critical for thymocyte development. Nature immunology *7*, 67-75.

Yang, R.B., Mark, M.R., Gray, A., Huang, A., Xie, M.H., Zhang, M., Goddard, A., Wood, W.I., Gurney, A.L., and Godowski, P.J. (1998). Toll-like receptor-2 mediates lipopolysaccharide-induced cellular signalling. Nature *395*, 284-288.

Zeng, X., Wei, Y.L., Huang, J., Newell, E.W., Yu, H., Kidd, B.A., Kuhns, M.S., Waters, R.W., Davis, M.M., Weaver, C.T., and Chien, Y.H. (2012). gammadelta T cells recognize a microbial encoded B cell antigen to initiate a rapid antigen-specific interleukin-17 response. Immunity *37*, 524-534.

Zhang, Y., Cado, D., Asarnow, D.M., Komori, T., Alt, F.W., Raulet, D.H., and Allison, J.P. (1995). The role of short homology repeats and TdT in generation of the invariant gamma delta antigen receptor repertoire in the fetal thymus. Immunity *3*, 439-447.

Zinkernagel, R.M., and Doherty, P.C. (1974). Restriction of in vitro T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. Nature *248*, 701-702.

APPENDIX I –

CYTOMEGALOVIRUS AND TUMOUR STRESS SURVEILLANCE BY BINDING OF A HUMAN γδ T CELL ANTIGEN RECEPTOR TO ENDOTHELIAL PROTEIN C RECEPTOR

Cytomegalovirus and tumor stress-surveillance by human γδ T cell receptor binding to Endothelial Protein C Receptor

Carrie R. Willcox, Vincent Pitard, Sonia Netzer, Lionel Couzi, Mahboob Salim, Tobias Silberzahn, Jean-François Moreau, Adrian C. Hayday, Benjamin E. Willcox, and Julie Déchanet-Merville

Nature Immunology: doi:10.1038/ni.2394

APPENDIX II –

COMPLETE CHEMICAL SHIFT DATA FOR SKINT IgV

Appendix II

Residue Number	Chemical Shift (ppm)												
	CD	CD1	CD2	CE	CE1	CE2	CE3	CG	CG1	CG2	CH2	CZ	CZ3
M1	-	-	-	-	-	-	-	-	-	-	-	-	-
S2	-	-	-	-	-	-	-	-	-	-	-	-	-
S3	-	-	-	-	-	-	-	-	-	-	-	-	-
E4	-	-	-	-	-	-	-	-	-	-	-	-	-
P5	50.08	-	-	-	-	-	-	27.46	-	-	-	-	-
F6	-	133.3	133.3	-	131.5	131.6	-	-	-	-	-	130.2	-
I7	-	14.04	-	-	-	-	-	-	26.9	18.6	-	-	-
V8	-	-	-	-	-	-	-	-	24.46	22.22	-	-	-
N9	-	-	-	-	-	-	-	-	-	-	-	-	-
G10	-	-	-	-	-	-	-	-	-	-	-	-	-
L11	-	26.08	22.78	-	-	-	-	-	-	-	-	-	-
E12	-	-	-	-	-	-	-	36.4	-	-	-	-	-
G13	-	-	-	-	-	-	-	-	-	-	-	-	-
P14	50.52	-	-	-	-	-	-	25.3	-	-	-	-	-
V15	-	-	-	-	-	-	-	-	22.5	21.1	-	-	-
L16	-	25.04	23.86	-	-	-	-	27.78	-	-	-	-	-
A17	-	-	-	-	-	-	-	-	-	-	-	-	-
S18	-	-	-	-	-	-	-	-	-	-	-	-	-
L19	24.94	25.18	25.09	-	-	-	-	26.87	-	-	-	-	-
G20	-	-	-	-	-	-	-	-	-	-	-	-	-
G21	-	-	-	-	-	-	-	-	-	-	-	-	-
N22	-	-	-	-	-	-	-	-	-	-	-	-	-
L23	-	24.11	26.84	-	-	-	-	-	-	-	-	-	-
E24	-	-	-	-	-	-	-	37.14	-	-	-	-	-
L25	-	24.22	26.31	-	-	-	-	26.72	-	-	-	-	-
S26	-	-	-	-	-	-	-	-	-	-	-	-	-
C27	-	-	-	-	-	-	-	-	-	-	-	-	-
Q28	-	-	-	-	-	-	-	32.97	-	-	-	-	-
L29	-	23.01	26.25	-	-	-	-	25.52	-	-	-	-	-
S30	-	-	-	-	-	-	-	-	-	-	-	-	-
P31	48.95	-	-	-	-	-	-	26.46	-	-	-	-	-
P32	50.6	-	-	-	-	-	-	27.88	-	-	-	-	-
Q33	-	-	-	-	-	-	-	33.33	-	-	-	-	-
Q34	-	-	-	-	-	-	-	34.63	-	-	-	-	-
A35	-	-	-	-	-	-	-	-	-	-	-	-	-
Q36	-	-	-	-	-	-	-	-	-	-	-	-	-
H37	-	-	-	-	-	-	-	-	-	-	-	-	-
M38	-	-	-	-	-	-	-	-	-	-	-	-	-
E39	-	-	-	-	-	-	-	37.98	-	-	-	-	-

I40	-	26.23	-	-	-	-	-	-	14.75	15.9	-	-	-
R41	42.91	-	-	-	-	-	-	28.33	-	-	-	-	-
W42	-	127.9	-	-	-	-	120.1	-	-	-	122.4	-	122.5
F43	-	132.2	132.2	-	129.8	129.8	-	-	-	-	-	128.4	-
L53	-	24.05	26.15	-	-	-	-	27.86	-	-	-	-	-
Y54	-	133.9	133.8	-	117.8	117.8	-	-	-	-	-	-	-
R55	43.9	-	-	-	-	-	-	27.67	-	-	-	-	-
D56	-	-	-	-	-	-	-	-	-	-	-	-	-
G57	-	-	-	-	-	-	-	-	-	-	-	-	-
K58	29.08	-	-	42.14	-	-	-	24.01	-	-	-	-	-
D59	-	-	-	-	-	-	-	-	-	-	-	-	-
M60	-	-	-	-	-	-	-	-	-	-	-	-	-
F61	-	-	-	-	-	-	-	-	-	-	-	-	-
G62	-	-	-	-	-	-	-	-	-	-	-	-	-
E63	-	-	-	-	-	-	-	36.78	-	-	-	-	-
I64	-	13.78	-	-	-	-	-	-	-	17.57	-	-	-
I65	-	13.98	-	-	-	-	-	-	26.5	18.34	-	-	-
S66	-	-	-	-	-	-	-	-	-	-	-	-	-
K67	29.31	-	-	42.14	-	-	-	24.91	-	-	-	-	-
Y68	-	132.6	132.6	-	119.7	119.7	-	-	-	-	-	-	-
E70	-	-	-	-	-	-	-	-	-	-	-	-	-
R71	-	-	-	-	-	-	-	-	-	-	-	-	-
T72	-	-	-	-	-	-	-	-	-	22.88	-	-	-
E73	-	-	-	-	-	-	-	35.96	-	-	-	-	-
L74	-	26.08	24.21	-	-	-	-	26.96	-	-	-	-	-
L75	-	23.17	24.53	-	-	-	-	27.44	-	-	-	-	-
K76	29.24	-	-	42.43	-	-	-	24.69	-	-	-	-	-
D77	-	-	-	-	-	-	-	-	-	-	-	-	-
G78	-	-	-	-	-	-	-	-	-	-	-	-	-
I79	-	14.39	-	-	-	-	-	-	31.75	17.45	-	-	-
G80	-	-	-	-	-	-	-	-	-	-	-	-	-
E81	-	-	-	-	-	-	-	36.67	-	-	-	-	-
G82	-	-	-	-	-	-	-	-	-	-	-	-	-
K83	29.77	-	-	42.3	-	-	-	25.36	-	-	-	-	-
V84	-	-	-	-	-	-	-	-	19.36	22.59	-	-	-
T85	-	-	-	-	-	-	-	-	-	20.77	-	-	-
L86	-	25.25	26.06	-	-	-	-	23.51	-	-	-	-	-
R87	44.3	-	-	-	-	-	-	27.07	-	-	-	-	-
I88	-	14.64	-	-	-	-	-	-	27.33	17.55	-	-	-
F89	-	-	-	-	-	-	-	-	-	-	-	-	-
N90	-	-	-	-	-	-	-	-	-	-	-	-	-
V91	-	-	-	-	-	-	-	-	22.78	20.63	-	-	-
T92	-	-	-	-	-	-	-	-	-	21.44	-	-	-

V93	-	-	-	-	-	-	-	-	21.49	20.32	-	-	-
D94	-	-	-	-	-	-	-	-	-	-	-	-	-
D95	-	-	-	-	-	-	-	-	-	-	-	-	-
D96	-	-	-	-	-	-	-	-	-	-	-	-	-
G97	-	-	-	-	-	-	-	-	-	-	-	-	-
S98	-	-	-	-	-	-	-	-	-	-	-	-	-
Y99	-	132	132	-	117.2	117.2	-	-	-	-	-	-	-
H100	-	-	-	-	-	-	-	-	-	-	-	-	-
C101	-	-	-	-	-	-	-	-	-	-	-	-	-
V102	-	-	-	-	-	-	-	-	22.18	21.73	-	-	-
F103	-	133.1	133.1	-	132.9	132.9	-	-	-	-	-	-	-
K104	30.41	-	-	42.52	-	-	-	25.1	-	-	-	-	-
D105	-	-	-	-	-	-	-	-	-	-	-	-	-
G106	-	-	-	-	-	-	-	-	-	-	-	-	-
D107	-	-	-	-	-	-	-	-	-	-	-	-	-
F108	-	131.9	131.9	-	131.9	131.9	-	-	-	-	-	-	-
Y109	-	-	-	-	-	-	-	-	-	-	-	-	-
E110	-	-	-	-	-	-	-	37.55	-	-	-	-	-
E111	-	-	-	-	-	-	-	34.14	-	-	-	-	-
H112	-	-	-	-	-	-	-	-	-	-	-	-	-
I113	-	14.56	-	-	-	-	-	-	27.46	17.84	-	-	-
T114	-	-	-	-	-	-	-	-	-	21.91	-	-	-
E115	-	-	-	-	-	-	-	35.76	-	-	-	-	-
V116	-	-	-	-	-	-	-	-	22.23	24.31	-	-	-
K117	28.88	-	-	42.15	-	-	-	24.84	-	-	-	-	-
I118	-	12.3	-	-	-	-	-	-	28.27	18.03	-	-	-
T119	-	-	-	-	-	-	-	-	-	22.1	-	-	-
V121	-	-	-	-	-	-	-	-	-	-	-	-	-

Chemical shift data from all assignment experiments for the backbone and sidechain residues of Skint1 IgV